



XX WPI: 1997-394558/37.  
 DR Preparation of S-layer proteins by expressing sbs-A gene in Gram  
 PT negative bacterium - or new sbs-B gene in any host, also new  
 PT recombinant proteins containing heterologous inserts, e.g.  
 PT epitope(s), useful as vaccines and adjuvants  
 XX  
 PS Claim 1: Pages 9-14; 31pp; German.  
 CC The present sequence encodes the Bacillus stearothermophilus PV72  
 CC S-layer protein, sbs-A. S-layer structures can be used as vaccines  
 CC or adjuvants, particularly when they include a bacterial ghost that  
 CC may contain additional epitopes in its membrane. Other uses of  
 CC recombinant sbs-A, depending on the nature of the inserted peptide,  
 CC are as an universal carrier for biotinylated reactants for use in  
 CC immunological or hybridisation assays (the insert is streptavidin),  
 CC to induce immune responses (epitopes), as a reagent for removing  
 CC cytoline or toxin from serum (antigenic epitopes), as a molecular  
 CC spinning nozzle (polyhydroxybutyrate synthase) and as a molecular  
 CC laser (luciferase).  
 CC  
 XX Sequence 3687 BP: 1316 A; 660 C; 715 G; 996 T; 0 other;

Query Match 100.0%; Score 3687; DB 18; Length 3687;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggtatgagaaagagctgtgaactgaacagcagcgtcattgacgaagtcattt 60  
 Db 1 atggtatgagaaagagctgtgaactgaacagcagcgtcattgacgaagtcattt 60  
 QY 61 gtcgtcgcacaaacccaacgcttcgtgaagcgcgtacagatgtagcaacagtagtaagccaa 120  
 Db 61 gtcgtcgcacaaacccaacgcttcgtgaagcgcgtacagatgtagcaacagtagtaagccaa 120  
 QY 121 gcaaaagcagcttcaaaaagacatactactacagccatacagtaacggaacatggt 180  
 Db 121 gcaaaagcagcttcaaaaagacatactactacagccatacagtaacggaacatggt 180  
 QY 181 gaattcccaaaacataacagatgtatagtgaatacaacaagcggaaaaaagatacgt 240  
 Db 181 gaattcccaaaacataacagatgtatagtgaatacaacaagcggaaaaaagatacgt 240  
 QY 241 gatcgcgtagccttagtgataaagcaggtggcggaaaaaagacgcttacttagctgac 300  
 Db 241 gatcgcgtagccttagtgataaagcaggtggcggaaaaaagacgcttacttagctgac 300  
 QY 301 ttacaaaagaatctgaacttaacglttcaaaagcaaaccttaacatctggcggaagctcgt 360  
 Db 301 ttacaaaagaatctgaacttaacglttcaaaagcaaaccttaacatctggcggaagctcgt 360  
 QY 361 gtagaacttaacatcgaatgcttacaactatgcacaacaaattgaagcgaagtgccagaag 420  
 Db 361 gtagaacttaacatcgaatgcttacaactatgcacaacaaattgaagcgaagtgccagaag 420  
 QY 421 ctagaagcgtcgttcaacgcaaaagatttagaanaagcgaagaacatactacatacaaaatt 480  
 Db 421 ctagaagcgtcgttcaacgcaaaagatttagaanaagcgaagaacatactacatacaaaatt 480  
 QY 481 cctttagaataaactcgacagctcatttagatcggtatattgtaaaaacaaactcgt 540  
 Db 481 cctttagaataaactcgacagctcatttagatcggtatattgtaaaaacaaactcgt 540  
 QY 541 gattactctgctctacatcttaaaagcaagaacgaactcgcgacagcttaattat 600  
 Db 541 gattactctgctctacatcttaaaagcaagaacgaactcgcgacagcttaattat 600  
 QY 601 gatattacggttgcaatgaagcgcgcaagacgaagcgtgtgtaaaagcagcaattta 660  
 Db 601 gatattacggttgcaatgaagcgcgcaagacgaagcgtgtgtaaaagcagcaattta 660

QY 661 gacaaagctaaagctgctgttgatcaaatcaatacttaacaaagaagtaacagatgct 720  
 Db 661 gacaaagctaaagctgctgttgatcaaatcaatacttaacaaagaagtaacagatgct 720  
 QY 721 ttcaaaactgaacttaacgaagtagcgaaaaagattagatgacagatgaagctgcctt 780  
 Db 721 ttcaaaactgaacttaacgaagtagcgaaaaagattagatgacagatgaagctgcctt 780  
 QY 781 actcaaaagctgaagtagtgatgacttaacacatcaacaagaagctgttaataaca 840  
 Db 781 actcaaaagctgaagtagtgatgacttaacacatcaacaagaagctgttaataaca 840  
 QY 841 gcaatgaacgtagaagcgaacacataaattacacattcaagctgtgcgaatgaagataca 900  
 Db 841 gcaatgaacgtagaagcgaacacataaattacacattcaagctgtgcgaatgaagataca 900  
 QY 901 gtaaaagtaaacgtgtagtgcataaagtgagcgcgtgaacattccatttccttaatt 960  
 Db 901 gtaaaagtaaacgtgtagtgcataaagtgagcgcgtgaacattccatttccttaatt 960  
 QY 961 acggcagatgttcttctatctacagacggaacacatcaactgtgtagcttcaactcca 1020  
 Db 961 acggcagatgttcttctatctacagacggaacacatcaactgtgtagcttcaactcca 1020  
 QY 1021 ttcgaaataatatacggagtagataaagtagtagtaaggtattaaagacaaaatgycaaa 1080  
 Db 1021 ttcgaaataatatacggagtagataaagtagtagtaaggtattaaagacaaaatgycaaa 1080  
 QY 1081 gaatttaagaagaatgacatcttcacgttcgaatgtagcgtgtgtagtctcaagt 1140  
 Db 1081 gaatttaagaagaatgacatcttcacgttcgaatgtagcgtgtgtagtctcaagt 1140  
 QY 1141 ttgggaactaatgttaacaacaacacactctgtgaacttagcagcagtagcttcgacact 1200  
 Db 1141 ttgggaactaatgttaacaacaacacactctgtgaacttagcagcagtagcttcgacact 1200  
 QY 1201 gacgaacttaacagtagtagtattgataagtgttgaacacctgaaacctgttaaacagctcg 1260  
 Db 1201 gacgaacttaacagtagtagtattgataagtgttgaacacctgaaacctgttaaacagctcg 1260  
 QY 1261 aacgttactattacagatggttgaaactggaaaacgcatctccagtaattgcatctctt 1320  
 Db 1261 aacgttactattacagatggttgaaactggaaaacgcatctccagtaattgcatctctt 1320  
 QY 1321 ggttacaactatcattacaggtttaaagaagcgttgaatgaactgtgaacaataataact 1380  
 Db 1321 ggttacaactatcattacaggtttaaagaagcgttgaatgaactgtgaacaataataact 1380  
 QY 1381 gctatcaaatatgttaaaaacatactggtttacaatgcaagaagcttaacgagtagtctc 1440  
 Db 1381 gctatcaaatatgttaaaaacatactggtttacaatgcaagaagcttaacgagtagtctc 1440  
 QY 1441 actgcaaacgcatcagcccaactgttgtaacgcttctactactttggtgtgtcaact 1500  
 Db 1441 actgcaaacgcatcagcccaactgttgtaacgcttctactactttggtgtgtcaact 1500  
 QY 1501 ttatctacgtgtctcttacaacaacatgtttgggttaattggcgtggtgtgtgaatga 1560  
 Db 1501 ttatctacgtgtctcttacaacaacatgtttgggttaattggcgtggtgtgtgaatga 1560  
 QY 1561 gctggaacttatatcctggtcttcaatltcaacaacaacggttgcataaagttagacgaa 1620  
 Db 1561 gctggaacttatatcctggtcttcaatltcaacaacaacggttgcataaagttagacgaa 1620  
 QY 1621 tctacttagctgataactttgataatgtgaagaagaactgtgtacagttgttctct 1680  
 Db 1621 tctacttagctgataactttgataatgtgaagaagaactgtgtacagttgttctct 1680  
 QY 1681 gaactaaataatgtagcagcgttaaaatgttaacttaagtgccaaaagcgaagccttaa 1740  
 Db 1681 gaactaaataatgtagcagcgttaaaatgttaacttaagtgccaaaagcgaagccttaa 1740  
 QY 1741 gaaaataacatctatcaaatcaaaaattaaaaagcttgaagttccgataagtagttaa 1800

Db	1741	gaaataacacatccatcaaatccaataataaaaagcgtcgagctcgatlaaaggatttgaa	1800
OY	1801	tttaggcacggtttaaagagaaaaacatactgatttcaaaaactaaagacttaactgtcttaca	1860
Db	1801	tttaggcacggtttaaagagaaaaacatactgatttcaaaaactaaagacttaactgtcttaca	1860
OY	1861	gtctatagcgttaacgtcttaaaaaatggcgagcgtggaattaaagtaactgaaagctcaagaa	1920
Db	1861	gtctatagcgttaacgtcttaaaaaatggcgagcgtggaattaaagtaactgaaagctcaagaa	1920
OY	1921	ttctacgtggaagtctccagaagaatttaattacctttaatgctccaaacggtttcggtgagc	1980
Db	1921	ttctacgtggaagtctccagaagaatttaattacctttaatgctccaaacggtttcggtgagc	1980
OY	1981	acaatcacataagcgttcgaagtctgctgtagttaaaagcgggtgcgaactatctgtcttaca	2040
Db	1981	acaatcacataagcgttcgaagtctgctgtagttaaaagcgggtgcgaactatctgtcttaca	2040
OY	2041	gcgaagtgcacatacttcacagctagtgittgaagcgggttaactggttcacagatggacaacaaa	2100
Db	2041	gcgaagtgcacatacttcacagctagtgittgaagcgggttaactggttcacagatggacaacaaa	2100
OY	2101	gtgaaagttgcgtcgttaacccaatttagaagcgttaaccaaagggttacaatttagtagtctcggt	2160
Db	2101	gtgaaagttgcgtcgttaacccaatttagaagcgttaaccaaagggttacaatttagtagtctcggt	2160
OY	2161	aaaagctgcacaacgctcgttttaaaagattgctgcgaactgcgaatacttttagcaactaaat	2220
Db	2161	aaaagctgcacaacgctcgttttaaaagattgctgcgaactgcgaatacttttagcaactaaat	2220
OY	2221	atctatatacttaacaacttgaaagtgacagaacgttaacgaacaaacggtttaaaaaagttac	2280
Db	2221	atctatatacttaacaacttgaaagtgacagaacgttaacgaacaaacggtttaaaaaagttac	2280
OY	2281	aaaagttgattcttttaaaaagacgctgtagtcagtttaccacacttcgaaacgtttgtagcaggt	2340
Db	2281	aaaagttgattcttttaaaaagacgctgtagtcagtttaccacacttcgaaacgtttgtagcaggt	2340
OY	2341	caaaaaattcacatacccaattttagcgaagaattaaaaaactctagtggtttcttagtggt	2400
Db	2341	caaaaaattcacatacccaattttagcgaagaattaaaaaactctagtggtttcttagtggt	2400
OY	2401	ggcacaagttaactgtgcgaagaatttaacaacaacacggaatgggttaagtctggttactgcgaca	2460
Db	2401	ggcacaagttaactgtgcgaagaatttaacaacaacacggaatgggttaagtctggttactgcgaca	2460
OY	2461	actgtatacagttgcgtcccttaagacagatgccaatggttaaaatggtaacagctgctggttaca	2520
Db	2461	actgtatacagttgcgtcccttaagacagatgccaatggttaaaatggtaacagctgctggttaca	2520
OY	2521	tttaactggtgcttcgacaataacgcgaacaaagatgcgaanaatttggtcttggttaagtaaaatct	2580
Db	2521	tttaactggtgcttcgacaataacgcgaacaaagatgcgaanaatttggtcttggttaagtaaaatct	2580
OY	2581	tctactgtaagaaatbtagatgtagatggtgtaagttaacttaaggaaaaaagaattatttaatt	2640
Db	2581	tctactgtaagaaatbtagatgtagatggtgtaagttaacttaaggaaaaaagaattatttaatt	2640
OY	2641	cgtttacaacagcgtgcgagacaacactgttagcttctgtgaagctgctgtgcgaagaatggt	2700
Db	2641	cgtttacaacagcgtgcgagacaacactgttagcttctgtgaagctgctgtgcgaagaatggt	2700
OY	2701	caaaaacgcttcgcgtgcatttcccaaaaagacgcgaatgatatcaactaagaactta	2760
Db	2701	caaaaacgcttcgcgtgcatttcccaaaaagacgcgaatgatatcaactaagaactta	2760
OY	2761	gttgaattcaatgaaactgattatttagcggaaagttaaaactgcagaaacatcggttttaaagat	2820
Db	2761	gttgaattcaatgaaactgattatttagcggaaagttaaaactgcagaaacatcggttttaaagat	2820
OY	2821	gcagcaggttaatgctggttagcgtgtagctgtaacagcatlaagcgtttctacaaataaatt	2880
Db	2821	gcagcaggttaatgctggttagcgtgtagctgtaacagcatlaagcgtttctacaaataaatt	2880

Db	2821	gcacgacggttaatgcggtgaagctgtgtactgtgtacaacgcattgaagcggttctccaataattc	2880
Qy	2881	gtattcctccatctccaaagattaaaagctgtgtacagtttactctgttaacaattgacggt	2940
Db	2881	gtattcctccatctccaaagattaaaagctgtgtacagtttactctgttaacaattgacggt	2940
Qy	2941	gtgagagataagtaggttaacaacaatctctaatacattacttcgttcaagactgtatct	3000
Db	2941	gtgagagataagtaggttaacaacaatctctaatacattacttcgttcaagactgtatct	3000
Qy	3001	gcgagatccaaagttattcttcaacagatgtgcgacggtgcgacgttaagcgtttac	3060
Db	3001	gcgagatccaaagttattcttcaacagatgtgcgacggtgcgacgttaagcgtttac	3060
Qy	3061	aaaacaattaaattgtaattcagcgattcagttccaaacccaacaatcactcttaagaag	3120
Db	3061	aaaacaattaaattgtaattcagcgattcagttccaaacccaacaatcactcttaagaag	3120
Qy	3121	gcgtacgggaacctcattacttaataatcaactttagtaaatgtaataataaaca	3180
Db	3121	gcgtacgggaacctcattacttaataatcaactttagtaaatgtaataataaaca	3180
Qy	3181	tacaaaatgtattccacaaagggtgtgaacacctgtacgaggtttactctaataatgagtaga	3240
Db	3181	tacaaaatgtattccacaaagggtgtgaacacctgtacgaggtttactctaataatgagtaga	3240
Qy	3241	gttccaanaagattccaactcgtgtactgtatgtatgacaaagttaacattcacacaggt	3300
Db	3241	gttccaanaagattccaactcgtgtactgtatgtatgacaaagttaacattcacacaggt	3300
Qy	3301	tctgttgcctactgtacgaagaataaacctgcctctagttagcggttggtcatgysaatgysaaca	3360
Db	3301	tctgttgcctactgtacgaagaataaacctgcctctagttagcggttggtcatgysaatgysaaca	3360
Qy	3361	agctatactcaaggatgctgtcgaagacaacacgagctcggctcgttagcttgcgttcgagag	3420
Db	3361	agctatactcaaggatgctgtcgaagacaacacgagctcggctcgttagcttgcgttcgagag	3420
Qy	3421	ccagttgcctcctcaactctcagaaggatcagatttaacgaatgcgaactgtgcagatgaaca	3480
Db	3421	ccagttgcctcctcaactctcagaaggatcagatttaacgaatgcgaactgtgcagatgaaca	3480
Qy	3481	aataattactgtatgaataaacctgttgaagttattttccaagaagagatgtgtagacgacacat	3540
Db	3481	aataattactgtatgaataaacctgttgaagttattttccaagaagagatgtgtagacgacacat	3540
Qy	3541	gatgcaggtgctactaagaaggacattgaataaacaacaggttaactcctttagtacttgat	3600
Db	3541	gatgcaggtgctactaagaaggacattgaataaacaacaggttaactcctttagtacttgat	3600
Qy	3601	aacgcgaagaatttaagattgttgtgaagcggaggttaagaatgcagcagagtaattgtga	3660
Db	3601	aacgcgaagaatttaagattgttgtgaagcggaggttaagaatgcagcagagtaattgtga	3660
Qy	3661	gatactattacattctatattaaatga 3687	
Db	3661	gatactattacattctatattaaatga 3687	
RESULT	2	:	
X22748			
ID	X22748	standard; DNA; 3687 BP.	
XX			
AC	X22748:	?	
XX			
DT	24-AUG-1999	(first entry)	
XX			
DE	B	stearothermophilus sbSA DNA.	
XX			
KM	sbSA protein, S-layer protein; Gram-negative; prokaryotic host cell;		
KM	integration; cytoplasmic membrane; secretion; periplasmic space; toxin;		
KM	eukaryotic host cell; vaccine; adjuvant; immunogenic epitope; luciferase		
KM	immunostimulant; cytokine; polyhydroxybutyrate; PHB synthase; body fluid;		

KW	molecular laser; universal carrier molecule; monomolecular layer; ds.	
XX		
OS	Bacillus stearothermophilus.	
XX		
FH	Key	location/Qualifiers
FT	CDS	1..3687
FT		/*tag= a
FT		/product= "bsa"
FT	sig_peptide	1..90
FT		/*tag= b
FT	mat_peptide	91..3684
FT		/*tag= c
FX		
PN	DE19732829-A1.	
XX		
PD	04-FEB-1999.	
XX		
PE	30-JUL-1997; 97DE-1032829.	
XX		
PR	30-JUL-1997; 97DE-1032829.	
XX		
PA	(LUBI/) LUBITZ W.	
XX		
PI	Lubitz W, Resch S;	
XX		
DR	WPI: 1999-122189/11.	
DR	P-PSDB: W93252.	
XX		
PT	Producing S-layer proteins in Gram-negative bacteria or eukaryotes -	
PT	integrated into membranes or organelles or secreted into periplasma	
PT	or growth medium, and nucleic acid encoding S-layer proteins with	
PT	peptide insertions, used in vaccines or for enzymatic reactions	
XX		
PS	Claim 5; Page 9-14; 34pp; German.	

This invention describes a method for the production of a S-layer protein (I) which comprises (a) preparing a Gram-negative prokaryotic host cell transformed with nucleic acid (II) encoding (I), linked to a signal sequence (SS) that encodes a protein which causes at least one of (i) integration of (I) into the external or cytoplasmic membranes and/or (ii) secretion of (I) into the periplasmic space or extracellular medium, (b) culturing the cell to express (I) and (c) optionally recovering (I) from the membranes, periplasmic space and/or extracellular medium. Alternatively, a eukaryotic cell is used as host and then the SS, which is optional, promotes integration into the cytoplasmic membrane or an organelle and/or secretion into the extracellular medium. (I) and derived structures, may include a wide variety of polypeptide inserts and are useful as (i) vaccines or adjuvants (with immunogenic epitopes or immunostimulants inserts such as cytokines) (ii) as reactors (inserts are enzymes, e.g. polyhydroxybutyrate (PHB) synthase for use as a 'molecular spherule', for production of PHB or luciferase for use as molecular laser (when combined with substrate and oxygen) and (iii) as universal carrier molecule (streptavidin is inserted) for use in hybridisation and immuno assays, or for selective elimination of cytokines, toxins etc. from body fluids (inserts are specific binding epitopes). In this system, heterologous (I) do not form inclusion bodies but rather monomolecular layers, and in eukaryotic cells they undergo glycosylation. This sequence encodes the Bacillus stearothermophilus sbba protein which is used to illustrate the method of the invention.

Sequence 3687 BP; 1316 A; 660 C; 715 G; 996 T; 0 other;

Query Match	100.0%;	Score 3687;	DB 20;	Length 3687;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3687;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 atgagatagagaaaaagctgtgtgaaactagcaacacagcaagtgctatttcgacgaagtgcattt 60  
Db 1 atgagtagagaaaaaagctgtgtgaaactagcaacacagcaagtgctatttcgacgaagtgcattt 60  
QY 61 gtgcgctgcgaatcccaacagctcttcgaagcgggctacagatgttagcaacagtaagttaagccaa 120

Db	61	gfcgctgcgaactccaaacgcctctctgaaagcgggtctacagatgltgcaacagtagtlaagccaa	120
QY	121	gcaaaaagacagcttcacaaaagacataactacttacagccatacagtlaacggaataactg	180
Db	121	gcaaaaagacagcttcacaaaagacataactacttacagccatacagtlaacggaataactg	180
QY	181	gaattcccaacatlaacgagtatactgctgaatacacaaagcgaaaaaacgataacgct	240
Db	181	gaattcccaacatlaacgagtatactgctgaatacacaaagcgaaaaaacgataacgct	240
QY	241	gattggtgagatattagtagaataaagcaggtggtggtggtggtggtggtggtggtggtggtg	300
Db	301	ttacaaaaaagatatactgaacttcagcttcttcacaaagcgaacccataactctgcgaagctgc	360
QY	301	ttacaaaaaagatatactgaacttcagcttcttcacaaagcgaacccataactctgcgaagctgc	360
Db	301	ttacaaaaaagatatactgaacttcagcttcttcacaaagcgaacccataactctgcgaagctgc	360
QY	361	gtacgaactcttcacatgcgtcttcacacatgcaacaaatctagacgaaatgctgcgaagag	420
Db	361	gtacgaactcttcacatgcgtcttcacacatgcaacaaatctagacgaaatgctgcgaagag	420
QY	421	ctagaggtgtgtgtgtctcaagcaaaagattttgaaaaagcgaacacatcacatacacaatt	480
Db	421	ctagaggtgtgtgtgtctcaagcaaaagattttgaaaaagcgaacacatcacatacacaatt	480
QY	481	ccttcagaaatttaaaactctgcacagatcatttttagatctgcgttatgtgtaaacaactgct	540
Db	481	ccttcagaaatttaaaactctgcacagatcatttttagatctgcgttatgtgtaaacaactgct	540
QY	541	gattactctgcctctacattlaaagcaaaaagcgaagaaactctgcgaacgcttaattat	600
Db	541	gattactctgcctctacattlaaagcaaaaagcgaagaaactctgcgaacgcttaattat	600
QY	601	gattttccggttgcgaagaagcggcgcggaagatcacagaagcgtgtgaagcggcgcaatta	660
Db	601	gattttccggttgcgaagaagcggcgcggaagatcacagaagcgtgtgaagcggcgcaatta	660
QY	661	gacaaagcgtaaagctgt	720
Db	661	gacaaagcgtaaagctgt	720
QY	721	ttcaaaactgcactcaacagagtagcgaaaaaagcattagatgcagatgaaagctgcgtt	780
Db	721	ttcaaaactgcactcaacagagtagcgaaaaaagcattagatgcagatgaaagctgcgtt	780
QY	781	actcccaaaagtgtgaagtgtaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	840
Db	781	actcccaaaagtgtgaagtgtaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	840
QY	841	gcagctacccgtgaacggaacacacataaactttacacactttcagctgcgtgcgaatgaaatata	900
Db	841	gcagctacccgtgaacggaacacacataaactttacacactttcagctgcgtgcgaatgaaatata	900
QY	901	gtaaaagctaaatactgtgactatctataaagtgtgaacgtgaacatcccatcttgcccttaat	960
Db	901	gtaaaagctaaatactgtgactatctataaagtgtgaacgtgaacatcccatcttgcccttaat	960
QY	961	acggcgaagtgttctcttatctatacagcgggaaaaaactatactgcgtgtgattctcaactcca	1020
Db	961	acggcgaagtgttctcttatctatacagcgggaaaaaactatactgcgtgtgattctcaactcca	1020
QY	1021	ttcgaaaaataatcgggtgttaaaagtagttaaaggtattaaaggtattaaaggtattaaaggtatt	1080
Db	1021	ttcgaaaaataatcgggtgttaaaagtagttaaaggtattaaaggtattaaaggtattaaaggtatt	1080
QY	1081	gaatttaaaagagatgattcactcttcacaaagcttcgaatgtagtgcgtgtattctcaaatg	1140
Db	1081	gaatttaaaagagatgattcactcttcacaaagcttcgaatgtagtgcgtgtattctcaaatg	1140
QY	1141	tttggaaactaatgtataacaaacactctctgttaaaacttagacgagcaggtacttgcgaact	1200
Db	1141	tttggaaactaatgtataacaaacactctctgttaaaacttagacgagcaggtacttgcgaact	1200

QY	1201	gacgaacttcaacagtagtatlattgttaagttgtttagcacctgnaaactgtaaacagctcg	1260
Db	1201	gacgaacttcaacagtagtatlattgttaagttgtttagcacctgnaaactgtaaacagctcg	1260
QY	1261	aacgttaacttaacagatgtttgaaactggaanaacgcattccagttaattgcatactacttc	1320
Db	1261	aacgttaacttaacagatgtttgaaactggaanaacgcattccagttaattgcatactacttc	1320
QY	1321	ggcttcaacattactctttagcgtttaaagaagcgttttagtaactgtttaacaataaact	1380
Db	1321	ggcttcaacattactctttagcgtttaaagaagcgttttagtaactgtttaacaataaact	1380
QY	1381	gctatcaataatggtttaaacaattaaactggtttcaaatgcagaagaacttaacgattagttc	1440
Db	1381	gctatcaataatggtttaaacaattaaactggtttcaaatgcagaagaacttaacgattagttc	1440
QY	1441	actgcaaacgcgcacgcaccaactgtttgctaaccgtctactacttaagttgtaacaact	1500
Db	1441	actgcaaacgcgcacgcaccaactgtttgctaaccgtctactacttaagttgtaacaact	1500
QY	1501	ttatctacacggtttctcttccaacaanaatgtttggggttaaatgtgcgttggttgtaaga	1560
Db	1501	ttatctacacggtttctcttccaacaanaatgtttggggttaaatgtgcgttggttgtaaga	1560
QY	1561	gctggaacttatatctctgtctctcaattcacacaacagctttgctactaagtttagacga	1620
Db	1561	gctggaacttatatctctgtctctcaattcacacaacagctttgctactaagtttagacga	1620
QY	1621	tctactttaagctgataactttggtatagtgtgaanaaagaaactggtacagttgtgtcttc	1680
Db	1621	tctactttaagctgataactttggtatagtgtgaanaaagaaactggtacagttgtgtcttc	1680
QY	1681	gaacctaanaataatgcaagtcagaagcttaaatggttaactttagttgccaagaagggaccta	1740
Db	1681	gaacctaanaataatgcaagtcagaagcttaaatggttaactttagttgccaagaagggaccta	1740
QY	1741	gaanaatacatctatccaagaacttaanaaagaagcgtttgaagtcggaataaagttatgaa	1800
Db	1741	gaanaatacatctatccaagaacttaanaaagaagcgtttgaagtcggaataaagttatgaa	1800
QY	1801	ttaggaactgtttaacgagaanaacatltgagttccaanaactcaagaacttaactgtctctaca	1860
Db	1801	ttaggaactgtttaacgagaanaacatltgagttccaanaactcaagaacttaactgtctctaca	1860
QY	1861	gttatttagtgataactgtcttaanaatggcgcagcgcgttgattaaagttaactgtaagctcaaga	1920
Db	1861	gttatttagtgataactgtcttaanaatggcgcagcgcgttgattaaagttaactgtaagctcaaga	1920
QY	1921	tttactgtgaagttcttcagagaatttaataactttaatgcttcaacccggttttgggttagc	1980
Db	1921	tttactgtgaagttcttcagagaatttaataactttaatgcttcaacccggttttgggttagc	1980
QY	1981	acaatcacatacgtgtcaaatgttcgtgagttaaagaacgggtgccaacttatctgtcttaca	2040
Db	1981	acaatcacatacgtgtcaaatgttcgtgagttaaagaacgggtgccaacttatctgtcttaca	2040
QY	2041	gcaagtgcacaatctccagcttagtgtttgaagcgggttaacgtgccaagtgtggaacatacaca	2100
Db	2041	gcaagtgcacaatctccagcttagtgtttgaagcgggttaacgtgccaagtgtggaacatacaca	2100
QY	2101	gtgaagtggtgcgtcgaacaaattagaagcgttaaccagaaggtgcacaataatgagtggttcggt	2160
Db	2101	gtgaagtggtgcgtcgaacaaattagaagcgttaaccagaaggtgcacaataatgagtggttcggt	2160
QY	2161	aaaagtgcaaacgctctctgtttaaagatgtctgcacaatgccaatacttttagacataactat	2220
Db	2161	aaaagtgcaaacgctctctgtttaaagatgtctgcacaatgccaatacttttagacataactat	2220
QY	2221	atctatactttaacaactgaaaggttcaagagcttaacgacccaacggttacaagaatctc	2280
Db	2221	atctatactttaacaactgaaaggttcaagagcttaacgacccaacggttacaagaatctc	2280

OY	2261	aaagvgatctcttaaaagcgtctgacgttactactacactacgaaagttatgvcaggt	2340
Db	2281	aaaggtattctttaaagagcgcgtctgacgttactactacactacgaaagttatgvcaggt	2340
OY	2341	caaaatctacatcccaatttaacggaagattaaaactctcagtggtctcttggtgtgt	2400
Db	2341	caaaaatctcctacccaatttaacggaagattaaaactctcagtggtctcttggtgtgt	2400
OY	2401	ggcaaaatlaactgfcggaagatttaaaaacaaacggatggtgtatgctgcgtacgcgaa	2460
Db	2401	ggcaaaatlaactgfcggaagatttaaaaacaaacggatggtgtatgctgcgtacgcgaa	2460
OY	2461	actgtatcagtgctgccttaagacagatgtcaaatgtgtaaagtaacagcgtcgtcgtgtaca	2520
Db	2461	actgtatcagtgctgccttaagacagatgtcaaatgtgtaaagtaacagcgtcgtcgtgtaca	2520
OY	2521	ttaactgtctctgaacaataacggaagttgngaatttgcgtctgtgtgtatgaagct	2580
Db	2521	ttaactgtctctgaacaataacggaagttgngaatttgcgtctgtgtgtatgaagct	2580
OY	2581	tctactatcggaatttgcgtgtgtatgtagctgtgtatgttaatgaagaaagaaatcttaatt	2640
Db	2581	tctactatcggaatttgcgtgtgtatgtagctgtgtatgttaatgaagaaagaaatcttaatt	2640
OY	2641	cgttacaacacgctgagagacacactgtagcttctgtgaaagctcgtcgtcgcgacaagaatggt	2700
Db	2641	cgttacaacacgctgagagacacactgtagcttctgtgaaagctcgtcgtcgcgacaagaatggt	2700
OY	2701	caaaaagcctctgctgcatctcccaaaagcactgtgcaattgtatacaaacctaagagcttatta	2760
Db	2701	caaaaagcctctgctgcatctcccaaaagcactgtgcaattgtatacaaacctaagagcttatta	2760
OY	2761	gttgatctcaatgaagctgtatttagggggaagttaaaacccggaagacatcgtgttttaagat	2820
Db	2761	gttgatctcaatgaagctgtatttagggggaagttaaaacccggaagacatcgtgttttaagat	2820
OY	2821	gcagcaggtatgfcggtgtagctgtgtactgttaacagcatatgaacggttctacaataaatt	2880
Db	2821	gcagcaggtatgfcggtgtagctgtgtactgttaacagcatatgaacggttctacaataaatt	2880
OY	2881	gtattcactcctcactcaagatataaaagctgtgtacagtttactctcgttaacaattgcaggt	2940
Db	2881	gtattcactcctcactcaagatataaaagctgtgtacagtttactctcgttaacaattgcaggt	2940
OY	2941	gttgaggaataaagttaggttaacacaactcttaataatcattactcgtttaaagacgttact	3000
Db	2941	gttgaggaataaagttaggttaacacaactcttaataatcattactcgtttaaagacgttact	3000
OY	3001	gcgaatcccaacggttatcttcaatccagcatctgtgtacggtgcggttaaaagcttgacgttct	3060
Db	3001	gcgaatcccaacggttatcttcaatccagcatctgtgtacggtgcggttaaaagcttgacgttct	3060
OY	3061	aaaacaattcaattgaattcagcagcttcagtccaacccaacaacccaactcttaagaag	3120
Db	3061	aaaacaattcaattgaattcagcagcttcagtccaacccaacaacccaactcttaagaag	3120
OY	3121	gcttagcgaacttcaatttactactaatttcaactttagtaaatgttaataatgaaataaaaaa	3180
Db	3121	gcttagcgaacttcaatttactactaatttcaactttagtaaatgttaataatgaaataaaaaa	3180
OY	3181	tacaaaattgatatcccaaaagggtgtataacacttgcaggtttactacatatagtattgaca	3240
Db	3181	tacaaaattgatatcccaaaagggtgtataacacttgcaggtttactacatatagtattgaca	3240
OY	3241	gtttcaaaagatttccaacgtgtactgtatattgatagcaaaagttaacatcatcacaaggt	3300
Db	3241	gtttcaaaagatttccaacgtgtactgtatattgatagcaaaagttaacatcatcacaaggt	3300
OY	3301	tctgttctcctctgaggaagtaaaaactgtctcagtagagcgtgtgtgtctatgtgaaatggaca	3360
Db	3301	tctgttctcctctgaggaagtaaaaactgtctcagtagagcgtgtgtgtgtctatgtgaaatggaca	3360
OY	3361	agcatatctcaggaatgctgtgagcaacacagactcgtcgttatgcgaaattcgttgcgag	3420

```
Db 3361 agctatactcagatgctgcagcacacagactcgcgtctgtagcttcgttcgag 3420
|||||
QY 3421 ccagttgccccttaactcagaagagatgatttaacgaatgcaactgagcaatgaca 3480
|||||
Db 3421 ccagttgccccttaactcagaagagatgatttaacgaatgcaactgagcaatgaca 3480
|||||
QY 3481 aatactcagatgataaactggttgaagttatttcaaaaagaagtcgacgacacat 3540
|||||
Db 3481 aatactcagatgataaactggttgaagttatttcaaaaagaagtcgacgacacat 3540
|||||
QY 3541 gatgcaggtgctcactcaagagacattagtaataacacagttactccttagtactgtat 3600
|||||
Db 3541 gatgcaggtgctcactcaagagacattagtaataacacagttactccttagtactgtat 3600
|||||
QY 3601 aacagcagaagctataagattgttgaagtgaggaagtaagaatgacaggtgaattgca 3660
|||||
Db 3601 aacagcagaagctataagattgttgaagtgaggaagtaagaatgacaggtgaattgca 3660
|||||
QY 3661 gatatactatcatctatataagtaa 3687
|||||
Db 3661 gatatactatcatctatataagtaa 3687
|||||

RESULT 3
T08695
ID T08695 standard: DNA; 3832 BP.
XX
AC T08695:
XX
DT 15-JUL-1996 (first entry)
XX
DE sbasA gene encoding S-layer protein.
XX
KM sbasA gene; S-layer; cell surface layer; expression; ss
XX
OS Bacillus stearothermophilus.
XX
FH Key Location/Qualifiers
FT CDS 37..3723
FT RBS 19..25 /*tag- a
FT 5'UTR 1..36 /*tag- b
FT 3'UTR 3723..3832 /*tag- c
FT sig_peptide 37..126 /*tag- d
FT mat_peptide 127..3720 /*tag- e
FT stem_loop 3769..3806 /*tag- f
FT /*tag- g /*note- "terminator"
XX
XX DE4425527-A1.
XX
XX 25-JAN-1996.
XX
XX 19-JUL-1994; 94DE-4425527.
XX
XX 19-JUL-1994; 94DE-4425527.
XX
XX (VOGE-) VOGELBUSCH GMBH.
XX
XX Lubitz W;
XX
XX WPI; 1996-077933/09.
XX
XX P-PSDB; R77673.
XX
XX Nucleic acid encoding signal peptide of Bacillus stearothermophilus
PT
```

```
PT S-layer protein - which has a lysine content of at least 10 per
cent.
XX
XX PS Disclosure; Page 6-7; 12pp; German.
XX
XX CC The claimed signal peptide encoding sequence (see tag e of T08695)
CC is pref. operably linked at the 3' end to a protein encoding
CC sequence. The protein is pref. the S-layer protein (see tag f of
CC T08695). At the 5' end, the signal peptide encoding sequence is
CC pref. linked to an expression control sequence, pref. the sequence
CC given in T08696.
XX
XX SQ Sequence 3832 BP; 1355 A; 666 C; 749 G; 1042 T; 0 other;

Query Match 100.0%; Score 3687; DB 17; Length 3832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1 atgataaggaanaagctgtgaactcagcacagcaagtcatactgcagaatgcat 60
|||||
Db 37 atgataaggaanaagctgtgaactcagcacagcaagtcatactgcagaatgcat 96
|||||
QY 61 gtccgtgcaaatcccaaacgcttcgtgaagcgcgtacagatgtacaaacagtagtaagccaa 120
|||||
Db 97 gtccgtgcaaatcccaaacgcttcgtgaagcgcgtacagatgtacaaacagtagtaagccaa 156
|||||
QY 121 gcaaaagcagcttcaaaaagaatcattactactacagccatcagtaacggaatgtgt 180
|||||
Db 157 gcaaaagcagcttcaaaaagaatcattactactacagccatcagtaacggaatgtgt 216
|||||
QY 181 gaattcccaacacttaacagatgatatgtcgtgaatacaacaagaacgaaacacatacct 240
|||||
Db 217 gaattcccaacacttaacagatgatatgtcgtgaatacaacaagaacgaaacacatacct 276
|||||
QY 241 gatgcggtagcatatgaataaagaagtgcgcggaaaaaagcgttacttaagctgat 300
|||||
Db 277 gatgcggtagcatatgaataaagaagtgcgcggaaaaaagcgttacttaagctgat 336
|||||
QY 301 ttacaaaagaataatgaacttgcgtttccaagcaaacaccttaactctgcggaagctgt 360
|||||
Db 337 ttacaaaagaataatgaacttgcgtttccaagcaaacaccttaactctgcggaagctgt 396
|||||
QY 361 gtacgaacttaccatcgatgcttcaacatacgcacaaatlaagcgaatgcccagaag 420
|||||
Db 397 gtacgaacttaccatcgatgcttcaacatacgcacaaatlaagcgaatgcccagaag 456
|||||
QY 421 ctgaaggtgctgttcaagcaaaaagattagaaaaagcgaacaatlaacatacaaaatt 480
|||||
Db 457 ctgaaggtgctgttcaagcaaaaagattagaaaaagcgaacaatlaacatacaaaatt 516
|||||
QY 481 ccttatgaatataaactgcgcagatcatttttagatcgcgtatattgttaaaaactcgt 540
|||||
Db 517 ccttatgaatataaactgcgcagatcatttttagatcgcgtatattgttaaaaactcgt 576
|||||
QY 541 gattactcgcctcattataagcaaaaagcagaagactcgcgcagcgttaattat 600
|||||
Db 577 gattactcgcctcattataagcaaaaagcagaagactcgcgcagcgttaattat 636
|||||
QY 601 gatattacggttgcaatgaagaacgcgcgaagtacaaagcgtgtgaagaagcgaattta 660
|||||
Db 637 gatattacggttgcaatgaagaacgcgcgaagtacaaagcgtgtgaagaagcgaattta 696
|||||
QY 661 gacaagaactaaagctgctgttgcataatcaatcaattcttccaaaagttaagtgct 720
|||||
Db 697 gacaagaactaaagctgctgttgcataatcaatcaattcttccaaaagttaagtgct 756
|||||
QY 721 ttcaaaactgaactacagaagtagcgaaaaaagcattagatgcagaatgaagtcgctt 780
|||||
Db 757 ttcaaaactgaactacagaagtagcgaaaaaagcattagatgcagaatgaagtcgctt 816
|||||
QY 781 actcacaaggttgaagtgtaagtcgattacaactcaaaaacaaagctgttaattaca 840
|||||
```

Db 817 actccaaagtgtgaagtgtaagtgcgattacaacatccaaaaaagaagctgtgtaataca 876  
QY 841 gacgtaccatgtagaagaaacaaacttaacaacttcagctgctgtaaatgtgaataca 900  
Db 877 gcagtlacagtgaaacgaaacaaacttaacaacttcagctgctgtaaatgtgaataca 936  
QY 901 gtaaacgtaaactctgtaacgtatctataaagtgagcggtaaacatctcatttcgcttaac 960  
Db 937 gtaacgtaaactctgtaacgtatctataaagtgagcggtaaacatctcatttcgcttaac 996  
QY 961 acggcagatgtctcttattctatcacagcgaaatactacatctgtgagtgcttaacctca 1020  
Db 997 acggcagatgtctcttattctatcacagcgaaatactacatctgtgagtgcttaacctca 1056  
QY 1021 ttgaaataataacgagtagtaagtagtagtaagtagttaaggttaaggtcaaaaatggctaaa 1080  
Db 1057 ttgaaataataacgagtagtaagtagtagtaagtagttaaggttaaggtcaaaaatggctaaa 1116  
QY 1081 gaattaaagaagatgcatctcaacttcagcttcgaaatgatgctgtatctcaagtg 1140  
Db 1117 gaattaaagaagatgcatctcaacttcagcttcgaaatgatgctgtatctcaagtg 1176  
QY 1141 ttggaactaatgttaacaacaacacactctgttaaaccttagcagcaggtaacttcgacact 1200  
Db 1177 ttggaactaatgttaacaacaacacactctgttaaaccttagcagcaggtaacttcgacact 1236  
QY 1201 gaagatacttaacagatgatttgtaagttgtttagcactgtgaactgttaaaagctcg 1260  
Db 1237 gaagatacttaacagatgatttgtaagttgtttagcactgtgaactgttaaaagctcg 1296  
QY 1261 aacgttactttaacagatgatttgtaagttgtttagcactgtgaactgttaaaagctcg 1320  
Db 1297 aacgttactttaacagatgatttgtaagttgtttagcactgtgaactgttaaaagctcg 1356  
QY 1321 ggttctacaacttaactatctacgtgttaaaagaagcgtttagtaactgtgttaacaataaactt 1380  
Db 1357 ggttctacaacttaactatctacgtgttaaaagaagcgtttagtaactgtgttaacaataaactt 1416  
QY 1381 gctatacatatgtttaaaacaacttaactgtttacaatgacagaagcttcgagtttagtttc 1440  
Db 1417 gctatacatatgtttaaaacaacttaactgtttacaatgacagaagcttcgagtttagtttc 1476  
QY 1441 acgtgaaacgcatacagcacaacactgttgcatacgcgtctactactttaggtgtgacact 1500  
Db 1477 acgtgaaacgcatacagcacaacactgttgcatacgcgtctactactttaggtgtgacact 1536  
QY 1501 ttactactgttctctcttacaacaacatgtttgggttaaatgtgcgtggtgggtgtaata 1560  
Db 1537 ttactactgttctctcttacaacaacatgtttgggttaaatgtgcgtggtgggtgtaata 1596  
QY 1561 gctggaactatatactctgttcttccaattcaacaacagcttgcatacgaattgaacgaa 1620  
Db 1597 gctggaactatatactctgttcttccaattcaacaacagcttgcatacgaattgaacgaa 1656  
QY 1621 tctactttagctgtaaacctgttatagttgaaaaagaatctgtgtacagttgtgtctct 1680  
Db 1657 tctactttagctgtaaacctgttatagttgaaaaagaatctgtgtacagttgtgtctct 1716  
QY 1681 gaactaaataataatgacagacgcttaaaatgttaactttagtgcacaaagcggaccttaaa 1740  
Db 1717 gaactaaataataatgacagacgcttaaaatgttaactttagtgcacaaagcggaccttaaa 1776  
QY 1741 gaaataataatctatacacaatttaaaaagcgttgaagtcggttaagaaggtattgaa 1800  
Db 1777 gaaataataatctatacacaatttaaaaagcgttgaagtcggttaagaaggtattgaa 1836  
QY 1801 ttgagcactgttaacggaacacatagatgttcaaaaactcaagaacttaactgtcttaca 1860  
Db 1837 ttgagcactgttaacggaacacatagatgttcaaaaactcaagaacttaactgtcttaca 1896  
QY 1861 gttatttagcgttaacggttaaaaatgtgcagcgttggatttaaaagttaacgtgaagctcaagaa 1920  
Db 1897 gttatttagcgttaacggttaaaaatgtgcagcgttggatttaaaagttaacgtgaagctcaagaa 1956

QY 1921 ttactgtgaagttctcaagaatttaaatatacttaattgctacaacccgttccggtagc 1980  
Db 1957 ttactgtgaagttctcaagaatttaaatatacttaattgctacaacccgttccggtagc 2016  
QY 1981 acatacacatacagtcgaagttgtgtgtagtaaaagcgggtgcacaactatctgcttaca 2040  
Db 2017 acatacacatacagtcgaagttgtgtgtagtaaaagcgggtgcacaactatctgcttaca 2076  
QY 2041 gcaagtgacatactccagctagttgttgaagcgggttaactgtgtcaagaatgaaacatacaa 2100  
Db 2077 gcaagtgacatactccagctagttgttgaagcgggttaactgtgtcaagaatgaaacatacaa 2136  
QY 2101 gtgaagttgtgtgttaacccaattagaacgtaaaccaaggggtacaaaatttaggtgttcggt 2160  
Db 2137 gtgaagttgtgtgttaacccaattagaacgtaaaccaaggggtacaaaatttaggtgttcggt 2196  
QY 2161 aaagtgtaacagatccctgtttaaagaatgtctgcacaatgcacaactcttagcaactaactat 2220  
Db 2197 aaagtgtaacagatccctgtttaaagaatgtctgcacaatgcacaactcttagcaactaactat 2256  
QY 2221 atctatacatctcaactgtgaaggtcaagacgttaacagcacaacgggttacaaaagtattc 2280  
Db 2257 atctatacatctcaactgtgaaggtcaagacgttaacagcacaacgggttacaaaagtattc 2316  
QY 2281 aaagtgattctttaaagaacgctgtgaatgcaggttactaacttcgaaacgttgaatgcaggt 2340  
Db 2317 aaagtgattctttaaagaacgctgtgaatgcaggttactaacttcgaaacgttgaatgcaggt 2376  
QY 2341 caaaaattcacatcccaatttaacgaagaattaaacactcttaggtgttctttagtggt 2400  
Db 2377 caaaaattcacatcccaatttaacgaagaattaaacactcttaggtgttctttagtggt 2436  
QY 2401 ggcgaagtaacgttgcgaagaatttaacaacaacgggtatgggtagaatgtgttactgtgaaca 2460  
Db 2437 ggcgaagtaacgttgcgaagaatttaacaacaacgggtatgggtagaatgtgttactgtgaaca 2496  
QY 2461 actgtacagattgtctcttaagaacagatgcgaatggtgaagaatgaacgctgtgtgttaca 2520  
Db 2497 actgtacagattgtctcttaagaacagatgcgaatggtgaagaatgaacgctgtgtgttaca 2556  
QY 2521 ttaactgtcttgaacaatacgaacaaagatgcgaatgtgctgtgtgtagtagtaagttct 2580  
Db 2557 ttaactgtcttgaacaatacgaacaaagatgcgaatgtgctgtgtgtagtagtaagttct 2616  
QY 2581 tctactgtgaagtgtgtgtgtatgtatgtgtatgttaattaaaggaataattttaact 2640  
Db 2617 tctactgtgaagtgtgtgtgtatgtatgtgtatgttaattaaaggaataattttaact 2676  
QY 2641 cgtttacaacagcgtggagacacacactgtatgcttctgttgaagcgtctgtgcacaaagtgt 2700  
Db 2677 cgtttacaacagcgtggagacacacactgtatgcttctgttgaagcgtctgtgcacaaagtgt 2736  
QY 2701 caaaacgcttctgtgtatcccaacaacagcgtcaatttatatacgaatgaagcttata 2760  
Db 2737 caaaacgcttctgtgtatcccaacaacagcgtcaatttatatacgaatgaagcttata 2796  
QY 2761 gttgaattcaagaaactgtattagcggaaagttaaacctgtgaacacatcgttgttaagat 2820  
Db 2797 gttgaattcaagaaactgtattagcggaaagttaaacctgtgaacacatcgttgttaagat 2856  
QY 2821 gcagcaggtatgtcgttagtgtagtctgttaacagattagacggttcttaaaaataatt 2880  
Db 2857 gcagcaggtatgtcgttagtgtagtctgttaacagattagacggttcttaaaaataatt 2916  
QY 2881 gttactcactcactcaagaatcaaaaagcgtgtacagtttactcctgttaacaattgtcaggt 2940  
Db 2917 gttactcactcactcaagaatcaaaaagcgtgtacagtttactcctgttaacaattgtcaggt 2976  
QY 2941 gtgaagataaagtagtaacacacatctctaaatacattactcgttcaagaactgtatct 3000  
Db 2977 gtgaagataaagtagtaacacacatctctaaatacattactcgttcaagaactgtatct 3036



QY	3001	gcgatcccaagcttattctttaacacgcatctgtcgaaggcgcaagttaacogtttcgcacctct	3060
Db	3037	ggcaatcccaagcttattctttaacacgcatctgtcgaaggcgcaagttaacogtttcgcacctct	3096
OY	3061	aaacaattacaatttgattcacgagtttcagttcccaaccaccaatcactctttaagaag	3120
Db	3097	aaacaattacaatttgattcacgagtttcagttcccaaccaccaatcactctttaagaag	3156
OY	3121	gctgacggaacttcatttacttaataatlacactttagtaaatgtaaatgnaatataaaaca	3180
Db	3157	gctgacggaacttcatttacttaataatlacactttagtaaatgtaaatgnaatataaaaca	3216
OY	3181	tacaaaaattgtattcccacaagaagtgtaaaccttgacaggaatttaccataatgagttaca	3240
Db	3217	tacaaaaattgtattcccacaagaagtgtaaaccttgacaggaatttaccataatgagttaca	3276
OY	3241	gttccaaaagaatttccaactgtgtactcagatattgatagcaaaagttacattccatccacagt	3300
Db	3277	gttccaaaagaatttccaactgtgtactcagatattgatagcaaaagttacattccatccacagt	3336
OY	3301	tcgtgtgctaactgcgaagaagtaaaacctgctcctagtagagcgttgtgtcatgnaatgaaaca	3360
Db	3337	tcgtgtgctaactgcgaagaagtaaaacctgctcctagtagagcgttgtgtcatgnaatgaaaca	3396
OY	3361	agctatactcaagatgctgcgaacacacgacttcggtcgtctgtagctgacttcgttcgcgag	3420
Db	3397	agctatactcaagatgctgcgaacacacgacttcggtcgtctgtagctgacttcgttcgcgag	3456
OY	3421	cgaattgcccttcaattctcagaaggatcagatttatcgaatgcgaactgtgcgaattaaca	3480
Db	3457	cgaattgcccttcaattctcagaaggatcagatttatcgaatgcgaactgtgcgaattaaca	3516
OY	3481	aatttactgatataaaaactgtgtgaagttattttccaagaagagagtgtgacgcgaacacat	3540
Db	3517	aatttactgatataaaaactgtgtgaagttattttccaagaagagagtgtgacgcgaacacat	3576
OY	3541	gatvcagagtgctactaaaggagacattagtaataaacacagctactcctttagtactgcat	3600
Db	3577	gatvcagagtgctactaaaggagacattagtaataaacacagctactcctttagtactgcat	3636
OY	3601	aacgcgaagagctttaaagattgtgttaagtggaaggttaagaatgcagcgatattgttca	3660
Db	3637	aacgcgaagagctttaaagattgtgttaagtggaaggttaagaatgcagcgatattgttca	3696
OY	3661	gatactatcatctatataatgaataa 3687	
Db	3697	gatactatcatctatataatgaataa 3723	
<b>RESULT 4</b>			
X78246	ID	X78246 standard; DNA; 3768 BP.	
XX	AC	X78246;	
XX	D7	24-AUG-1999 (first entry)	
DE	XX	B. stearothermophilus sbxA/Bacteriophage fd Gene 3 fusion protein DNA.	
XX	XX	sbsA protein; S-layer protein; Gram-negative; prokaryotic host cell;	
KW	KW	integration; cytoplasmic membrane; secretion; periplasmic space; toxin;	
KW	KW	eukaryotic host cell; vaccine; adjuvant; immunogenic epitope; luciferase;	
KW	KW	immunostimulant; cytokine; polyhydroxybutyrate; PHB synthase; body fluid;	
KM	KM	molecular laser; universal carrier molecule; monomolecular layer; Gene 3;	
KM	fusion gene; ds.		
XX	OS	Synthetic.	
OS	OS	Bacteriophage fd.	
OS	OS	Bacillus stearothermophilus	
XX	XX	DE19732829 -A1.	
XX	XX	04-FEB-1999.	
DD	DD		

XX 30-JUL-1997: 97DE-1032829.  
PE  
XX  
PR 30-JUL-1997: 97DE-1032829.  
XX  
PA (LUBI/) LUBITZ W.  
XX  
PI Lubitz W, Resch S;  
DR WPI; 1999-122189/11.  
XX  
XX  
PT Producing S-layer proteins in Gram-negative bacteria or eukaryotes -  
PT integrated into membranes or organelles or secreted into periplasm  
PT or growth medium, and nucleic acid encoding S-layer proteins with  
PT peptide insertions, used in vaccines or for enzymatic reactions  
PS  
XX  
XX Example 8; Fig 5; 34pp; German.  
XX  
CC This invention describes a method for the production of a S-layer protein  
CC (I) which comprises (a) preparing a Gram-negative prokaryotic host cell  
CC transformed with nucleic acid (II) encoding (I), linked to a signal  
CC sequence (SS) that encodes a protein which causes at least one of (i)  
CC integration of (I) into the external or cytoplasmic membranes and/or  
CC (ii) secretion of (I) into the periplasmic space or extracellular medium  
CC (b) culturing the cell to express (I) and (c) optionally recovering (I)  
CC from the membranes, periplasmic space and/or extracellular medium.  
CC Alternatively, a eukaryotic cell is used as host and then the SS, which  
CC is optional, promotes integration into the cytoplasmic membrane or an  
CC organelle and/or secretion into the extracellular medium. (I), and  
CC derived structures, may include a wide variety of polypeptide inserts and  
CC are useful as (i) vaccines or adjuvants (with immunogenic epitopes or  
CC immunostimulants inserts such as cytokines) (ii) as reactants (inserts  
CC are enzymes, e.g. polyhydroxybutyrate (PHB) synthase for use as a  
CC 'molecular spinnerette' for production of PHB or luciferase for use as  
CC molecular laser (when combined with substrate and oxygen) and (iii) as  
CC universal carrier molecule (streptavidin is inserted) for use in  
CC hybridisation and immuno assays, or for selective elimination of  
CC cytokines, toxins etc. from body fluids (inserts are specific binding  
CC epitopes). In this system, heterologous (I) do not form inclusion bodies  
CC but rather monomolecular layers, and in eukaryotic cells they undergo  
CC glycosylation. This sequence encodes a fusion gene constructed from the  
CC Bacillus stearothermophilus sbsA gene and the bacteriophage fd Gene 3  
CC signal sequence which is used in the method of the invention.  
CC  
XX  
XX  
SQ Sequence 3768 BF; 1320 A; 683 C; 733 G; 1032 T; 0 other;  
SQ

Query Match 97.68; Score 3597.4; DB 20; Length 3768;  
Best Local Similarity 99.38; Pred. No. 0;  
Matches 3613; Conservative 0; Mismatches 26; Indels 0; Gaps 0

0Y 49 gcaagtgcattgtgcgtgtcaaatcccaacacgcttcgaagcggctacagatgtagcaaca 108  
DBb 19 gcaattcccttaagtgtgttccttctctatagcggccagcgcgctcagatgtagcaaca 78  
0Y 109 gtatgaagcccaagcaaaagacacagttcaaaaaagcactactatacttagcgcatacagta 168  
DB 79 gtatgaagcccaagcaaaagacacagttcaaaaaagcactactatacttagcgcatacagta 138  
0Y 169 acggaactggtgaattcccaacattacagatgtatatgctgaatacaacaagaagcgaa 228  
DB 139 acggaactggtgaattcccaacattacagatgtatatgctgaatacaacaagaagaa 198  
0Y 229 aaacgataccgtgtatggtgtagcattagtgtaataaagcgggtgcccggaaaaaagcgct 288  
DB 199 aaacgataccgtgtatggtgtagcattagtgtaataaagcgggtgcccggaaaaaagcgct 258  
0Y 289 tacttagctgtattcaaaaaagaatatgaactacagtttccaagcaaacccctaactc 348  
DB 259 tacttagctgtattcaaaaaagaatatgaactacagtttccaagcaaacccctaactc 318  
0Y 349 ggcgaaactcgtgtagcaacttatacatgcagtgttacaactatgcaacaaatlagacga 408



Db 319 ggcgaagctcgttgaagcaactacatcgatcgtcttaacaactatgcaacaaattagacgaa 378  
QY 409 atggcgcaaaagcgtaggggtgctgcttcaagcaaaaagatttagaaaaagcagaacatcac 468  
Db 379 atggcgcaaaagcgttaggggtgctgcttcaagcaaaaagatttagaaaaagcagaacatcac 438  
QY 469 tatcaacaaatccctta tgaataatcaaacctgcgcacagctcatctttatagatcggtatggt 528  
Db 439 tatcaacaaatcccttcttgtaataataaacctgcgcacagctcatctttatagatcggtatggt 498  
QY 529 aaaaacactcgtgatttacttcgcgtcatatlttaagcaaaagcacaagaacttcgcgac 588  
Db 499 aaaaacactcgtgatttacttcgcgtcatatlttaagcaaaagcacaagaacttcgcgac 558  
QY 589 agttaattatgatattatattaccgtttgcaatgaaagcgcgcgaaggtacacagacgcgtgaaa 648  
Db 559 agcttaattatgatattatattaccgtttgcaatgaaagcgcgcgaaggtacacagacgcgtgaaa 618  
QY 649 gcagcgcaatttagacaagaagctgaagctgctgtgatcaatccaatctacttaccacaaa 708  
Db 619 gcagcgcaatttagacaagaagctgaagctgctgtgatcaatccaatctacttaccacaaa 678  
QY 709 gtaacagatgctttcaaaaactgtaacatacagaagtagcgaaaaagcatttagatgcagat 768  
Db 679 gtaacagatgctttcaaaaactgtaacatacagaagtagcgaaaaagcatttagatgcagat 738  
QY 769 gaagcgcgcgttacttccaaaagtgtgaaggtgaaggtcgatatacactccaacaaagct 828  
Db 739 gaagcgcgcgttacttccaaaagtgtgaaggtgaaggtcgatatacactccaacaaagct 798  
QY 829 gttgaattacacagcagatcagcgtgaagcgaacactaaattacaacttttagctgcgcga 888  
Db 799 gttgaattacacagcagatcagcgtgaagcgaacactaaattacaacttttagctgcgcga 858  
QY 889 aatgaagatacagtaaaacytaaaactgtaactgtaactataaagtgtgcgcgttaacatcca 948  
Db 859 aatgaagatacagtaaaacytaaaactgtaactgtaactataaagtgtgcgcgttaacatcca 918  
QY 949 ttggcccttaatacgcgaagttgtttcttcttaccagaacggaaaaactatacactgtgcgat 1008  
Db 919 ttggcccttaatacgcgaagttgtttcttcttaccagaacggaaaaactatacactgtgcgat 978  
QY 1009 gcttcaactccatcgcgaataataacagagatataaagtagttaaggtattaaagac 1068  
Db 979 gcttcaactccatcgcgaataataacagagatataaagtagttaaggtattaaagac 1038  
QY 1069 aaaaatgycaaaagaatttaaaagaagatgcatctcaacttcaagcttcgaaatgactgta 1128  
Db 1039 aaaaatgycaaaagaatttaaaagaagatgcatctcaacttcaagcttcgaaatgactgta 1098  
QY 1129 gtaaccctaaagtgtttggaactaatgttaacaaacacactctgttaaaacttagcagcaggt 1188  
Db 1099 gtaaccctaaagtgtttggaactaatgttaacaaacacactctgttaaaacttagcagcaggt 1158  
QY 1189 acttctgcacactgaagatcttaacagtagtattgtataagttgttagcaccggaact 1248  
Db 1159 acttctgcacactgaagatcttaacagtagtattgtataagttgttagcaccggaact 1218  
QY 1249 gtaaacagctcgaaagcttactatcatagatgttgaanaactggaanaagcacttccagtaatc 1308  
Db 1219 gtaaacagctcgaaagcttactatcatagatgttgaanaactggaanaagcacttccagtaatc 1278  
QY 1309 gcatctactctcgtgttcaacattactatacgtttaaagaagcgttagtaactgtgtaaa 1368  
Db 1279 gcatctactctcgtgttcaacattactatacgtttaaagaagcgttagtaactgtgtaaa 1338  
QY 1369 caataataaacttgctatcaataatgttaaaacatttaacttgtaacatgtagaagcttac 1428  
Db 1339 caataataaacttgctatcaataatgttaaaacatttaacttgtaacatgtagaagcttac 1398  
QY 1429 gaattaggtgttcactgcanaacgcatacagcaccactgtgtcaccgcttactactactta 1488  
Db 1399 gaattaggtgttcactgcanaacgcatacagcaccactgtgtcaccgcttactactactta 1458  
QY 1489 ggtgtgaacttatctactgttctcttacaacaaatgttgggggtaaatgtgcgtgt 1548  
Db 1459 ggtgtgaacttatctactgttctcttacaacaaatgttgggggtaaatgtgcgtgt 1518  
QY 1549 ggtgtgaatgaagctggaactttaaactgtgtcttcaacttcaacaaagcttgcgttact 1608  
Db 1519 ggtgtgaatgaagctggaactttaaactgtgtcttcaacttcaacaaagcttgcgttact 1578  
QY 1609 aagttagaagaacttactttagctgataaacttgcatttagtgaanaaagaactcgtgtaca 1668  
Db 1579 aagttagaagaacttactttagctgataaacttgcatttagtgaanaaagaactcgtgtaca 1638  
QY 1669 gttgtgtctctgaacataataataatgtgcgaagcgttaaaatgttaacttagtgcgcaaaa 1728  
Db 1639 gttgtgtctctgaacataataataatgtgcgaagcgttaaaatgttaacttagtgcgcaaaa 1698  
QY 1729 gcggaaccttaagaaaaataacatctatcaaatcaaaaatttaaaaagcgttgaagctcgat 1788  
Db 1699 gcggaaccttaagaaaaataacatctatcaaatcaaaaatttaaaaagcgttgaagctcgat 1758  
QY 1789 aaggtatgtaattagcgcactgttgaacgagaaaaacatatgagttcaaaaactcaagactta 1848  
Db 1759 aaggtatgtaattagcgcactgttgaacgagaaaaacatatgagttcaaaaactcaagactta 1818  
QY 1849 actgcttctcaatattatagcgttaagcttcaaaaatgycgcgcgtgattaaagtaact 1908  
Db 1819 actgcttctcaatattatagcgttcaaaaatgycgcgcgtgattaaagtaact 1878  
QY 1909 gaagctcaagaatttactgtggaagttctcagagaatttaaatcatlttaatgtgtacaacc 1968  
Db 1879 gaagctcaagaatttactgtggaagttctcagagaatttaaatcatlttaatgtgtacaacc 1938  
QY 1969 gtttcgggtatgacaacatcacatagcgttcaagttcgtgtatgaanaagcgcgtgtgcaactta 2028  
Db 1939 gtttcgggtatgacaacatcacatagcgttcaagttcgtgtatgaanaagcgcgtgtgcaactta 1998  
QY 2029 tctgcctcttaacagaagtgacatcatctccagctgaaggtgtgaagcgttactgttcaagat 2089  
Db 1999 tctgcctcttaacagaagtgacatcatctccagctgaaggtgtgaagcgttactgttcaagat 2058  
QY 2089 ggaacatacaaaagtgaaggtgtgtctgaacaaattgaagaaggtacaaggtgtacaactta 2148  
Db 2059 ggaacatacaaaagtgaaggtgtgtctgaacaaattgaagaaggtacaaggtgtacaactta 2118  
QY 2149 gtagtgttcggttaaaagtgacaagctcctgttaaaagatgctgcgaatgcaaatctttaa 2208  
Db 2119 gtagtgttcggttaaaagtgacaagctcctgttaaaagatgctgcgaatgcaaatctttaa 2178  
QY 2209 gcaactaactatatactataataatgaactgaaggtcaagaagcgttaacgaccaaagcgtt 2268  
Db 2179 gcaactaactatatactataataatgaactgaaggtcaagaagcgtttaacgaccaaagcgtt 2238  
QY 2269 acaaaagatctcaaaagtgatcttctttaaagaagcgtgcgtacgtttaaacttaacgaac 2326  
Db 2239 acaaaagatctcaaaagtgatcttctttaaagaagcgtgcgtacgtttaaacttaacgaac 2298  
QY 2329 gttgagtgcaggttcaaaaatcaactatccaatttagcgaagaatttaaaaacttctagtgt 2388  
Db 2299 gttgagtgcaggttcaaaaatcaactatccaatttagcgaagaatttaaaaacttctagtgt 2358  
QY 2389 tcttttagtgggtgcgcaaaatgaactgtcgaagaaatttaacaaacaggaatggttagatgtc 2448  
Db 2359 tcttttagtgggtgcgcaaaatgaactgtcgaagaaatttaacaaacaggaatggttagatgtc 2418  
QY 2449 ggtacgcggaacaaactgtatcaggtgtctcctaagaacagatgcaaatgttaagaatcagact 2508  
Db 2419 ggtacgcggaacaaactgtatcaggtgtctcctaagaacagatgcaaatgttaagaatcagact 2478  
QY 2509 gctgtgtttacattaaactgtgtcttgcnaataagcaataaagatgcgaagaattgcgtctgtga 2568  
Db 2479 gctgtgtttacattaaactgtgtcttgcnaataaagcaataaagatgcgaagaattgcgtctgtga 2538

QY	2569	gtagaagaagctctctactcgaagagattcgcgagctgctgacgctgctgaatgctaaattgaagaaaaa	2628
Db	2559	gtagaataagctctctactgattgattgattgcgcgagctgctgctgaatgctaaattgaagaaaaa	2598
QY	2629	gatatatttaattctgttcaacaagctgtgagacacacactgtgactctgtgtgaagaagctgcgt	2688
Db	2599	gatatatttaattctgttcaacaagctgtgagacacacactgtgactctgtgtgaagaagctgcgt	2658
QY	2689	gacaaagaatgtgtcaaaaacgcgtctgcgcgtgacatctcccaagaagcagctgcgaattgatacaact	2748
Db	2659	gacaaagaatgtgtcaaaaacgcgtctgcgcgtgacatctcccaagaagcagctgcgaattgatacaact	2718
QY	2749	aagagctattatgattgtgaattcaattcaatgaaaacggaatttagcggaaagttaaactgtagaacatc	2808
Db	2719	aagagctattatgattgtgaattcaattcaatgaaaacggaatttagcggaaagttaaactgtagaacatc	2778
QY	2809	gtctgttaagaagctgcgaacgagtaatgcggtgtgcgtgtgactgtgaacagcatatagaacggtctc	2868
Db	2779	gtctgttaagaagctgcgaacgagtaatgcggtgtgcgtgtgactgtgaacagcatatagaacggtctc	2838
QY	2869	acaaataaattgtgatccactccacatcccaagaatttaaaagctgtgacgtttacactgtta	2928
Db	2839	acaaataaattgtgatccactccacatcccaagaatttaaaagctgtgacgtttacactgtta	2898
QY	2929	acaaattgcaggtgtgtgagagataaagtatgattgaacacaaactcttaaaatacatatcctgttc	2988
Db	2899	acaaattgcaggtgtgtgagagataaagtatgattgaacacaaactcttaaaatacatatcctgttc	2958
QY	2989	aagagctgtatctgtgaaatcccaagcttatcttcaataagaattgtgcgtgcgaagctgtacatc	3048
Db	2959	aagagctgtatctgtgaaatcccaagcttatcttcaataagaattgtgcgtgcgaagctgtacatc	3018
QY	3049	gttgcacggttctaaacaacatatacaattgaaattgaagcgattcttcacaaaccacaacatc	3108
Db	3019	gttgcacggttctaaacaacatatacaattgaaattgaagcgattcttcacaaaccacaacatc	3078
QY	3109	actcttaagaagcggtgtgcgcgaacactcttaactacatatacctttgttaattgtataaat	3168
Db	3079	actcttaagaagcggtgtgcgcgaacactcttaactacatatacctttgttaattgtataaat	3138
QY	3169	gaaataataaacaatacaaaatgttatctccacaagcggtatacaactgtgacagagtttctcaa	3228
Db	3139	gaaataataaacaatacaaaatgttatctccacaagcggtatacaactgtgacagagtttctcaa	3198
QY	3229	tatgagttaagcagttctccaagaagatttccaacactgtgtactgcgatattgtatgacaaagtata	3288
Db	3199	tatgagttaagcagttctccaagaagatttccaacactgtgtactgcgatattgtatgacaaagtata	3258
QY	3289	tctatccagaagttctgtgtgttactgtgcgaagataaaacccgcctctgtatggcggttggctca	3348
Db	3259	tctatccagaagttctgtgtgttactgtgcgaagataaaacccgcctctgtatggcggttggctca	3318
QY	3349	tgtgaattgtgaanaaagctatactcagatgtcgtgcagcaacagacttcoggtcgtgtagctgac	3408
Db	3319	tgtgaattgtgaanaaagctatactcagatgtcgtgcagcaacagacttcoggtcgtgtagctgac	3378
QY	3409	tctcgttcgcgagccagttgtgcgcctctcaattctcagaagagatcgtatcttaagaagatgacaact	3468
Db	3379	tctcgttcgcgagccagttgtgcgcctctcaattctcagaagagatcgtatcttaagaagatgacaact	3438
QY	3469	gtgcgcagctaaacaatactatctatctgtatgtataaaacactgcttgaaagtattcttccaaaagagtgta	3528
Db	3439	gtgcgcagctaaacaatactatctatctgtatgtataaaacactgcttgaaagtattcttccaaaagagtgta	3498
QY	3529	gagcgacgaaccaatgtatgcaggtgtgctactaaagagacaattgataattacaacagttactcct	3588
Db	3499	gagcgacgaaccaatgtatgcaggtgtgctactaaagagacaattgataattacaacagttactcct	3558
QY	3589	ttagttagattgataacagcaagagctttaaagattgtttgtatagctgggggtttaaagatgagcgaca	3648
Db	3559	ttagttagattgataacagcaagagctttaaagattgtttgtatagctgggggtttaaagatgagcgaca	3618
QY	3649	ggtatagttgtcagataactatatactctatataagtaa	3687

Db 3619 ggaatgttgagatactattcaattcattataagtaa 3657  
|||||  
RESULT 5  
ID X78245  
XX X78245 standard; DNA; 4988 BP.  
AC X78245.  
DT 24-AUG-1999 (first entry)  
DE B. stearothermophilus sbSA/male fusion protein DNA sequence.  
KW sbSA protein; S-layer protein; Gram-negative; prokaryotic host cell;  
KW integration; cytoplasmic membrane; secretion; periplasmic space; toxin;  
KW eukaryotic host cell; vaccine; adjuvant; immunogenic epitope; luciferase;  
KW immunostimulant; cytokine; polyhydroxybutyrate; PHB synthase; body fluid;  
KW molecular laser; universal carrier molecule; monomolecular layer;  
KW fusion protein; male; ss.  
XX  
XX Synthetic.  
OS Bacillus stearothermophilus.  
XX  
XX DE19732829-A1.  
XX  
XX 04-FEB-1999.  
XX  
XX 30-JUL-1997; 97DE-1032829.  
XX  
XX 30-JUL-1997; 97DE-1032829.  
XX  
XX (LUBI./) LUBITZ W.  
PA Lubitz W, Resch S;  
PI  
XX  
XX WPI: 1999-122189/11.  
XX  
XX Producing S-layer proteins in Gram-negative bacteria or eukaryotes -  
PT integrated into membranes or organelles or secreted into periplasma  
PR or growth medium, and nucleic acid encoding S-layer proteins with  
PT peptide insertions, used in vaccines or for enzymatic reactions  
PS  
XX Example 7; Fig 4; 34pp; German.  
XX  
XX This invention describes a method for the production of a S-layer protein  
CC (I) which comprises (a) preparing a Gram-negative prokaryotic host cell  
CC transformed with nucleic acid (II) encoding (I), linked to a signal  
CC sequence (SS) that encodes a protein which causes at least one of (1)  
CC integration of (I) into the external or cytoplasmic membranes and/or  
CC (11) secretion of (I) into the periplasmic space or extracellular medium,  
CC (11) culturing the cell to express (I) and (c) optionally recovering (I)  
CC from the membranes, periplasmic space and/or extracellular medium.  
CC Alternatively, a eukaryotic cell is used as host and then the SS, which  
CC is optional, promotes integration into the cytoplasmic membrane or an  
CC organelle and/or secretion into the extracellular medium. (I), and  
CC derived structures, may include a wide variety of polypeptide inserts and  
CC are useful as (i) vaccines or adjuvants (with immunogenic epitopes or  
CC immunostimulants inserts such as cytokines) (ii) as reactors (inserts  
CC are enzymes, e.g. polyhydroxybutyrate (PHB) synthase for use as a  
CC 'molecular spinnetette' for production of PHB or luciferase for use as  
CC molecular laser (when combined with substrate and oxygen)) and (iii) as  
CC universal carrier molecule (streptavidin is inserted) for use in  
CC hybridisation and immuno assays, or for selective elimination of  
CC cytokines, toxins etc. from body fluids (inserts are specific binding  
CC epitopes). In this system, heterologous (I) do not form inclusion bodies  
CC but rather monomolecular layers, and in eukaryotic cells they undergo  
CC glycosylation. This sequence encodes a novel fusion protein constructed  
CC from the Bacillus stearothermophilus sbSA gene and the male gene signal  
CC sequence which is used to illustrate the method of the invention.  
XX  
XX Sequence 4988 BP; 1678 A; 985 C; 1052 G; 1273 T; 0 other;  
XX



Db 3349 gatgctgcaaatgcaaaactttagcgaactacatactacatcttaacatgaaagt 3408  
 QY 2245 caagaagctaacgacacacacaggtttacaaaagtctcaaaagtatcttttaaaagagct 2304  
 Db 3409 caagacgttaacacacacacaggtttacaaaagtctcaaaagtatcttttaaaagagct 3468  
 QY 2305 gtagcagttactacacttaagaaagctgtgacggttcaaaaatctacatccaaattagc 2364  
 Db 3469 gtagcagttactacacttaagaaagctgtgacggttcaaaaatctacatccaaattagc 3528  
 QY 2365 gaagaattaaaaactcttagtcttcttagtgggtgacgaagtacgttcgagaaatt 2424  
 Db 3529 gaagaattaaaaactcttagtcttcttagtgggtgacgaagtacgttcgagaaatt 3588  
 QY 2425 acaaaacagcgtggtgtagtctgctgacgtggaacacgttatacgttctgctcaagaca 2484  
 Db 3589 acaaaacagcgtggtgtagtctgctgacgtggaacacgttatacgttctgctcaagaca 3648  
 QY 2485 gatgcaaatgttaaaagtaaacagcgtctgtgttaacatctgctctgacataacagac 2544  
 Db 3649 gatgcaaatgttaaaagtaaacagcgtctgtgttaacatctgctctgacataacagac 3708  
 QY 2545 aaagatgacgaagtgcgtctgtgtatgataagctctctctcgtatggaattgctgtatga 2604  
 Db 3709 aaagatgacgaagtgcgtctgtgtatgataagctctctctcgtatggaattgctgtatga 3768  
 QY 2605 gctggtatgttaatttaaggaagaaagataatttaattctgtaacaaagcgttgagacact 2664  
 Db 3769 gctggtatgttaatttaaggaagaaagataatttaattctgtaacaaagcgttgagacact 3828  
 QY 2665 gtagctctgtgaaagctgtctgtgacaaagatgtgtcaaaagcgttctgtgcatccca 2724  
 Db 3829 gtagctctgtgaaagctgtctgtgacaaagatgtgtcaaaagcgttctgtgcatccca 3888  
 QY 2725 acaagcactgcaattgtatacaactaaagcttattagttcaattcaattgaactgattta 2784  
 Db 3889 acaagcactgcaattgtatacaactaaagcttattagttcaattcaattgaactgattta 3948  
 QY 2785 ggcgaagttaaaccttgagacacatggtgttaaaagatgcagcaggttaatgctgagctggt 2844  
 Db 3949 ggcgaagttaaaccttgagacacatggtgttaaaagatgcagcaggttaatgctgagctggt 4008  
 QY 2845 actgttaacagcatagacggtttctacaataattgttatccatccatcccaagaatta 2904  
 Db 4009 actgttaacagcatagacggtttctacaataattgttatccatccatcccaagaatta 4068  
 QY 2905 aaagctgtgacagtttctctctgttaacaactgacggtgtgagagataagtagtaacaca 2964  
 Db 4069 aaagctgtgacagtttctctctgttaacaactgacggtgtgagagataagtagtaacaca 4128  
 QY 2965 atctctaataaagttacttctgtcaagacgttatctgcgaatcccaagttatctcaatc 3024  
 Db 4129 atctctaataaagttacttctgtcaagacgttatctgcgaatcccaagttatctcaatc 4188  
 QY 3025 agcattgtgacggttgacgttaacggttgacggttcaaaaatctacatctgaatcagc 3084  
 Db 4189 agcattgtgacggttgacgttaacggttgacggttcaaaaatctacatctgaatcagc 4248  
 QY 3085 gattcagttccaaaacccaacatcactcttaagaaagctgacggaactcattactaat 3144  
 Db 4249 gattcagttccaaaacccaacatcactcttaagaaagctgacggaactcattactaat 4308  
 QY 3145 tacaactttaaataatgttaataatgaataataaacaataaataatgttatcccaaaagt 3204  
 Db 4309 tacaactttaaataatgttaataatgaataataaacaataaataatgttatcccaaaagt 4368  
 QY 3205 gtaacaactgacgagttactacataatagatagcagttccaaaagatttcaaacaggt 3264  
 Db 4369 gtaacaactgacgagttactacataatagatagcagttccaaaagatttcaaacaggt 4428  
 QY 3265 actgatatgtatgacgaagttaacatcacaaggtctgttctactgacgaagtaaaa 3324

Db 4429 actgatatgtatgacgaagttaacatcacaaggtctgttctactgacgaagtaaaa 4488  
 QY 3325 cctgcctatgaagcgttggttcatgtgaatggaacaagctatactcagaatgctgacga 3384  
 Db 4489 cctgcctatgaagcgttggttcatgtgaatggaacaagctatactcagaatgctgacga 4548  
 QY 3385 acaagcattgctgtcgtacgtgacgttcgttgcgagacaggtgacctcaatccagaa 3444  
 Db 4549 acaagcattgctgtcgtacgtgacgttcgttgcgagacaggtgacctcaatccagaa 4608  
 QY 3445 ggtatgatttaacgaatgcaactgtgacgtggaacaaataatctatgataaactgtt 3504  
 Db 4609 ggtatgatttaacgaatgcaactgtgacgtggaacaaataatctatgataaactgtt 4668  
 QY 3505 gaagttattccaagaagagtagacgacgacacatgacgtgacgttctactaagagaca 3564  
 Db 4669 gaagttattccaagaagagtagacgacgacacatgacgtgacgttctactaagagaca 4728  
 QY 3565 ttagtaattacaacagttactcctttagtactgtatgaacagcaagactataagattgtt 3624  
 Db 4729 ttagtaattacaacagttactcctttagtactgtatgaacagcaagactataagattgtt 4788  
 QY 3625 gtaagtggagttaaagatgcagcaggtatgtgacagatattctattctattaaag 3684  
 Db 4789 gtaagtggagttaaagatgcagcaggtatgtgacagatattctattctattaaag 4848  
 QY 3685 taa 3687  
 Db 4849 taa 4851

---

RESULT 6  
 ID T05868 standard; DNA: 3399 BP.  
 XX  
 AC T05868;  
 XX  
 DT 14-AUG-1996 (first entry)  
 XX  
 DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.  
 XX  
 KW Chicken leucocytozoan; Immunogen; recombinant vaccine; protection;  
 KW Immunisation; Vaccination; ss.  
 XX  
 OS Chicken leucocytozoan.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..3399  
 FT /tag a  
 FT misc\_feature 1150..3218  
 FT /tag b  
 FT note "Fragment referred to in the claims, for  
 FT use as insert in a recombinant vaccine  
 FT against chicken leucocytozoan disease"  
 FT  
 XX JP07284392-A.  
 XX  
 XX PD 31-OCT-1995.  
 XX  
 XX PF 19-APR-1994; 94UP-0080643.  
 XX  
 XX PR 19-APR-1994; 94UP-0080643.  
 XX  
 PA (DOBU-) DOBUTSUO SEIBUNSHUGAKUENKI SEIZAI KYOKAI.  
 PA (KITA) KITASATO KENKUSHO SH.  
 XX  
 DR MPI; 1996-006311/01.  
 DR P-PSDB; R97866.  
 XX  
 PT Chicken leucocytozoan immunogenic protein - used in a recombinant  
 PT vaccine against chicken leucocytozoan disease  
 XX  
 PS Claim 6; Page 6-9; 35pp; Japanese.

XX T05868 encodes a chicken leucocytozoan immunogenic protein, this DNA  
 CC or a fragment of it can be used in a recombinant vaccine to immunise  
 CC against chicken leucocytozoan disease. The DNA is used in a vector  
 CC and operatively linked to an expression regulatory sequence as in  
 CC standard practice.

XX Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 other;

Query Match 2.1%; Score 77.4; DB 17; Length 3399;  
 Best Local Similarity 43.0%; Pred. No. 8.9e-08;  
 Matches 436; Conservative 0; Mismatches 576; Indels 3; Gaps 1;

QY 84 tgaagcggtacagtgtagcgaacgtagtaagccagcaagccagcttcaaaaaagc 143  
 DB 2064 tgaagaaagaagaagaagaagtaacacatagaataagaagaagatcgaagaagc 2123  
 QY 144 atactatacttaacgcatcagtaacggaacgtgtgaattcccaaatcattacagatgt 203  
 DB 2124 aatacatgtaagaagaagaagaagtaacacatgtaagaagaagaagaagagatga 2183  
 QY 204 atatgtctgaatacaacaagaacgaacacgataccgtgtgtagtgatgtgaataa 263  
 DB 2184 agaagtaatacactgaagaagaagaagaagtaacacatacgaagaagaagaagaagc 2243  
 QY 264 agcaggtgtagcgaagaagaagaagcgttacttagctgtttacaaaaagaatataccta 323  
 DB 2244 aacacatgtaagaagaagaagaagaagaatatacactgaagaagaagaagaagtaata 2303  
 QY 324 cgtttcaagaacacaccccaaatcttgccgaagctcgtgtagcaactatcagatctta 383  
 DB 2304 tgaagaagaagaagaagaagaagtaacacatgaagaagaagaagaagaagtaacata 2363  
 QY 384 caactatgcaacaacaaattagcgaagaatgcgcgaagaagctagaaggtgctgtcaagca 443  
 DB 2364 agaaaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2423  
 QY 444 agattagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 503  
 DB 2424 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 2483  
 QY 504 agtcaatttagatcggtatcggtatcggtatcggtatcggtatcggtatcggtatcg 563  
 DB 2484 aacacatgtaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2543  
 QY 564 agcaaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 623  
 DB 2544 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2603  
 QY 624 ggcggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 683  
 DB 2604 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2663  
 QY 684 tcaaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 743  
 DB 2664 agtaacacactgtaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2720  
 QY 744 agcgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 803  
 DB 2721 tgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2780  
 QY 804 tgcgttaacactcaaaaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 863  
 DB 2781 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2840  
 QY 864 aaaaatacaacttcaagctgctgcaaatgaagaacagtaaacgtaactgaactgaact 923  
 DB 2841 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2900  
 QY 924 ctataaagtgaagcgttaacacttcccttgaatacgcgcagatgtttcttatctac 983  
 DB 2901 aacacatgtaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2960

QY 984 agacgaaataactacactgtgtgattcgaactccattcgaataataatagagatataa 1043  
 DB 2961 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 3020  
 QY 1044 agtagtagttaaaggtatataagaacaaaatgycgaagaatttaagaagaatgca 1098  
 DB 3021 aaagtaacacactgtaagaagaagaagaagaagaagaagaagaagaagaagaagaatgta 3075

# RESULT 7

Q87587 standard; DNA; 1686 BP.

Q87587;

19-DEC-1995 (first entry)

DNA encoding Leucocytozoan protozoan structural protein epitope.

leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;

leucocytozoanosis; treatment; ss.

Leucocytozoan protozoa sp.

JP07089995-A.

04-APR-1995.

10-SEP-1993; 93JP-0226078.

10-SEP-1993; 93JP-0226078.

(DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SETZAI KYOKAI.

(NISS-) NISSEIKEN KK.

WPI; 1995-167252/22.

P-PSDB; R70491.

Immune inducing polypeptide against Leucocytozoan protozoan - useful

in production of vaccines for treatment of leucocytozoanosis in

fowl.

Claim 1: Page 12-14; 20pp; Japanese.

Q87587-89 encode polypeptides having a whole or partial epitope of a

structural protein of Leucocytozoan protozoan (see R70491-93). The

polypeptides and DNA encoding them are useful in the production of

vaccines for the treatment of leucocytozoanosis of fowl.

Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 other;

Query Match 1.7%; Score 62.4; DB 16; Length 1686;  
 Best Local Similarity 45.3%; Pred. No. 0.00013;  
 Matches 308; Conservative 0; Mismatches 366; Indels 6; Gaps 2;

QY 96 agatgtagcaaacagttagcgaagcaaacgacaggttcaaaaaagcgtactatccta 155  
 DB 141 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 200  
 QY 156 cagccatcacatgtaacggaacgtgtgaattcccaaacattacgatactatgctgaata 215  
 DB 201 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 260  
 QY 216 caacaaagcgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 275  
 DB 261 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 320  
 QY 276 gaanaagagcgttacttgaattacaaaaagaatatacgaactgaactgaactgaact 335  
 DB 321 agaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 380



KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;  
XX infection; diagnosis; characterisation; detection; ds.  
OS Borrelia burgdorferi.  
XX MO9858943-A1.  
XX 30-DEC-1998.  
XX 18-JUN-1998: 98MO-US12764.  
XX 03-SEP-1997: 97US-0057483.  
XX 20-JUN-1997: 97US-0050359.  
XX 22-JUL-1997: 97US-0053344.  
XX 22-JUL-1997: 97US-0053377.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (MED-) MEDIMUNE INC.  
PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;  
PI White OR;  
XX WPI: 1999-081217/07.  
XX New isolated Borrelia burgdorferi nucleic acids - used to develop  
PT products for the detection, diagnosis, characterisation, prevention  
PT and therapy of infections, particularly Lyme disease  
XX Claim 1: Page 994-996; 1128pp; English.  
XX X20248 to X20402 represent polynucleotide sequences isolated from  
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for  
CC the detection, diagnosis, characterisation, prevention and therapy of  
CC Bb infections, e.g. Lyme disease. They can also be used for the  
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs  
CC to a family of motile, spiral-shaped bacteria called Spirochetes.  
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and  
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as  
CC Lyme disease.  
XX Sequence 3653 BP; 1587 A; 348 C; 690 G; 1027 T; 1 other;  
XX

Query Match 1.6%; Score 57.8; DB 20; Length 3653;  
Best Local Similarity 49.8%; Pred. No. 0.0017;  
Matches 146; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 615 aatgaagcgcgcgaagtacagaagcgtgtaagcagcagcaattgacaaagctaaagc 674  
DB 1385 agtgaataaagtagaagatccgaagcttaagttgaggaagaagaataaacaagagaa 1444  
QY 675 tggctgtgataatcaataataacttaccataaagtaacagatgcttcaaaactgaact 734  
DB 1445 tacagaagaacgaataaataaataaagtaagtaacagaagaacaaacaaacgaa 1504  
QY 735 aacagaagtagcgaataaagcattagatgacagatgagctgcgtcttaccataaagtga 794  
DB 1505 agcagaacaaagaataaataaagaagagagagcaagaagaacaaataaagaagaaga 1564  
QY 795 aagtgaagtgcattacaactcaacaaagcgtgtgtaattacagcagctaccagtga 854  
DB 1565 gcaagaagaagaagcctaagcagaataaagaagcctaagaataaagaagaacaaacaa 1624  
QY 855 cgaataacataaattacaacttcagctgtgtaagtaagtaagtaagtaagtaagtaag 907  
DB 1625 agagaagaacaaacgaagaagcagaataaagaagaagaagaagaagaagaaga 1677

RESULT 10  
Z56910  
ID Z56910 standard; DNA: 4248 BP.  
XX Z56910;  
AC

XX 08-MAY-2000 (first entry)  
DT S. aureus RUSA266 strain DNA encoding antibiotic resistance proteins.  
XX  
DE Antibiotic resistance: methicillin; cytoplasmic peptidoglycan precursor;  
XX veterinary medicine; ds.  
XX Staphylococcus aureus.  
XX  
XX Key Location/Qualifiers  
XX CDS 154..1410  
XX /tag- a  
XX /note= "ORF418; specifically claimed for in claim 1"  
XX CDS 1497..3500  
XX /tag- b  
XX /note= "ORF667; specifically claimed for in claim 4"  
XX US6013507-A.  
XX 11-JAN-2000.  
XX 10-JUL-1996: 96US-0678614.  
XX 10-JUL-1995: 95US-0001045.  
XX (UYRQ ) UNIV ROCKEFELLER.  
XX Tomasz A, De Lencastre H;  
XX WPI: 2000-159876/14.  
XX New nucleic acid encoding staphylococcal antibiotic resistance  
PT proteins, used to identify agents that restore antibiotic sensitivity  
PT to bacteria, particularly methicillin-resistant Staphylococcus aureus  
XX  
XX Claims 1 and 4; Fig 11; 44pp; English.  
XX This represents a Staphylococcus aureus RUSA266 strain nucleotide  
CC sequence encoding proteins associated with antibiotic resistance. The  
CC proteins, which are essential for high level resistance to antibiotics,  
CC specifically methicillin, are enzymes involved in synthesis of  
CC cytoplasmic peptidoglycan precursors, so their inhibition will result in  
CC bacterial cell walls with altered mucopeptide content. The nucleic acid  
CC and the proteins are used for identification (or rational design) of  
CC agents, e.g. antisense nucleic acid or synthetic peptides, that  
CC antagonize the action of the proteins and can restore the sensitivity  
CC of bacteria to antibiotics against which they are normally resistant,  
CC useful in human or veterinary medicine. Inhibition of the proteins  
CC allows known antibiotics to be used effectively against resistant  
CC bacteria, obviating the need to develop new antibiotics.  
XX  
XX Sequence 4248 BP; 1790 A; 766 C; 743 G; 949 T; 0 other;  
XX

Query Match 1.5%; Score 56.8; DB 21; Length 4248;  
Best Local Similarity 41.6%; Pred. No. 0.0029;  
Matches 697; Conservative 0; Mismatches 942; Indels 35; Gaps 4;

QY 52 agtcatttgcgtgctgaatccaacgcgttcgaagcgcgcgtacagatgtagcaacagta 111  
DB 2544 aatgcagcctaagttcttctgtgttaagaacatggaagcgacaacgcgtcacaaattta 2603  
QY 112 gtaagcagaacaaagcacagttcaaaaagcattactactacagcgcatacagtaag 171  
DB 2604 gaaagagttaaaacgaagaatccaataattgaaatatactactagcttacgacaaca 2653  
QY 172 gaaactggtgattcccaacattacagatgatatgctggaatacaacaagaagaataa 231  
DB 2664 aaaaatgagtcct-----ataatgaagttcaaaacagctgcacacagctagaata 2711  
QY 232 cgataccgtgtagcgttagcatgtaataaagcagctggtgcggaataaagcgtttac 291



```

Db 2712 gctcaaatgctacagcttccaaatgcaacaatgaagaagtagcagaagctgtagcgca 2771
QY 292 ttgctgatttaccataaagaatatagaactagctttccaagaacacctaaatctgyc 351
Db 2772 gtaaatgagctcaaaagaggtttacatgacatccaagttgtttaaataccaagaa 2831
QY 352 gaactcgttgaacaacttccatcagatgctacacatgcaacaataatgaagaatg 411
Db 2832 gtgctgatacaaaatcaaaagatattagataaataatcaatgcaatccaacagcaaa 2891
QY 412 cgcacagagtagagcgctgctgttcaagcaaaagattagaaaaaagcagaactactat 471
Db 2892 gttaaacctgcagctgtagcgaagtagaanaagcatalataacgtaaacagaagaatt 2951
QY 472 cacaataatcctttagaataatgaacacgcagctcatatttagctgcgttagtgaataa 531
Db 2952 caaatacgaatgcttcaacacagaagaaaaaacaagctgcataacagattagaact 3011
QY 532 acaactcgtgatttacttcgtctacatttaagcaaaaagcacaagaacttcgcgacagc 591
Db 3012 aaaaagaagaagaagaacaacatctgtgctgcaaaatacaaacagtgatgtaacaa 3071
QY 592 ttaacttcatcatcaccgctgtgcaatgcaagcgcggaagtacagaagcgtctgtaagca 651
Db 3072 gctaaagacaatagatattgctgcaatlaatcaagttcagctgcacacactaagaatcg 3131
QY 652 ggcacatttagacaagaagtaaaagctgctgtgtgataaatacaatcacttccaaagta 711
Db 3132 g---atgcuaaagcggaatctgctcaaaaagcaagtgaaagctaaacacgcaattgaagca 3188
QY 712 acagatgcttcaaaaactgaaactaacagaagtagcgaaaaaagcattagatgcagatgaa 771
Db 3189 atgaaatgattcgtactctctgaagaacaacagcagcgaaagcaaaagtgatcaagcgta 3248
QY 772 gctgcttactccaagaagtgaagaatgtaagtgtagtgcattacactcaaaagctggtt 831
Db 3249 gtctctcaaacgctgtatatagtatgatactgtagcagaacaatgtagtgaatgcaaaa 3308
QY 832 gaattaacagcagtagcagtagaagcgaacataaatcactt-----cagctgctg 886
Db 3309 ctacaactgaagctacacatcgcagccatcaacctgtagcgaacatgcttaaacacagcaa 3368
QY 887 caaattgaaatagatacgtaaatcgtacgtatctatcaaatgtagcgtgatacttc 946
Db 3369 aacaaggaattgcagataaagtagacaagctcaagaagaacgcaattgtagaataacgct 3428
QY 947 caattgccttaatacggcagatgttcttcttatactacagcgaaaaaactacatctgtg 1006
Db 3429 caaacaactgaagaaaagcgctgtctaacaacaagaattcaaacctgaaaaaaacagctg 3488
QY 1007 atgcttgaactccatctggaataataatcggagtagtaagtagtgaagttaagttaag 1066
Db 3489 atgcgcgaatagatgacgacatacaataatgcggaagttagaagcggtctaaaaagcagcaa 3548
QY 1067 acaaaaatgccaagaatlttaaaagaagatgcattccacttccaagcttcogaatatgctg 1126
Db 3549 ttgctaaatggaagcgattcagccagcaacaacaactaaagataatgycgaagaagcaaa 3608
QY 1127 taattactcaagtgcttggaaactaatgttaaca-----acaacactctg 1171
Db 3609 ttgctacgaagcgaaatgaaagtaaaaacagcaatcgcctcaacgcaagacattactgctg 3668
QY 1172 taacttagcagcagtagtacttcgacactgcagctacttcaagtagtattgttagt 1231
Db 3669 aagaataatgcagcgctaaatgtagcgtagtaattgctgtgacacaagcaaatagcaca 3788
QY 1232 tgttagcactgaacacgttaaacagcgcgaagcttactattacagatgttgaactgaa 1291
Db 3729 ttgaagctgcttaatagtaacaatgtagtagcacaaggaagaacgacgggtgaaatgta 3788
QY 1292 aacgcatccagtaattgcatctactctgttacaataattactattacgttaagaag 1351

```

```

Db 3789 ttgatacgaatgaaccacagtttaataaaaagcaactgacgtlaatgaaatcacagaa 3848
QY 1352 cgttagtaactggttaacaataataacttgcattcaataatggttaaaacttaactgtt 1411
Db 3849 tttaataacaataatgcagaagattcaagtaagccagatgcacagatgaagaanaac 3908
QY 1412 acaatcagaagcttgaagttagtgcttcactgcaaacgcatcagcaccacgtgtgcta 1471
Db 3909 aagcagctgctgtagaagcaaatctgaaatgtgtaagcaaatcaagcatttcagcag 3968
QY 1472 cgcctcctactactttagtggtgtacaaacttattcactgtgttcttacaacaatgtt 1531
Db 3969 caacttactacgcacaaggttgatgaaagcaaatgcaagaagcagcatttaatgcg 4028
QY 1532 ggggttaattggtggtggtgtgaatgaagctgaacttattcctggtcttcaatca 1591
Db 4029 taacaccaaaagtgtgtaagaacaagcggttaaggttgaaattgataatcaaacaa 4088
QY 1592 caaacaagtttgcctactaagtttagcgaatcactactttagctgataacttgtttagt 1651
Db 4089 cgcaacaataatgtatcaataatgataatgatacgaagctacacagagaagaagaagcagta 4148
QY 1652 aaaaagaatcgtgtacagttgtgtgtcttgaactaaataataatgcagcagctta 1705
Db 4149 ttcaacaattgacacagcagttacagacgcgaaataataatattacagctgcaaa 4202

RESULT 11
X89891/c
ID X89891 standard; DNA; 397 BP.
XX
AC X89891;
XX
DT 05-NOV-1999 (first entry)
XX
DE Spino cerebellar ataxia type III (SCAIII) gene fragment.
XX
XX Spino cerebellar ataxia type III; SCAIII; reverse dot hybridisation;
KW PCR-microplate hybridisation; PCR-MPH; trinucleotide repeat; TNR;
KW SCAIII syndrome; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT repeat_region .137..355
FT FT /*tag= a
FT repeat_unit .137..139
FT FT /*tag= b
FT FT /note= "trinucleotide repeat"
PN MO9943852-A1.
XX
PD 02-SEP-1999.
XX
XX 18-FEB-1999; 99MO-KR00078.
XX
PR 26-FEB-1998; 98KR-0006278.
XX
PA (JIND/) JIN D K.
PA (SMSU ) SAMSUNG FINE CHEM CO LTD.
PI Jin DK;
XX
XX WPI; 1999-527634/44.
XX
XX Diagnosis of spino cerebellar ataxia type III (SCA III) syndrome
PT using techniques which ensure highly accurate diagnosis
XX
XX Claim 1; Page 12-13; 28pp; English.
XX
XX The invention relates to the diagnosis of spino cerebellar ataxia type
CC III (SCAIII) syndrome using reverse dot hybridisation or PCR-microplate
CC hybridisation (PCR-MPH). The method comprises attaching a portion of the

```

CC SCAIII gene containing 73 copies of the trinucleotide (CAG) repeat unit  
 CC (the present sequence) to a substrate, and hybridizing with amplified  
 CC testee genomic DNA containing copies of the trinucleotide units. PCR  
 CC amplified with labeled primers (X89889-90). The new method is useful for  
 CC diagnosis of SCAIII syndrome, and for determining the severity of the  
 CC disease. The present sequence represents the SCAIII gene fragment  
 CC containing 73 trinucleotide (TNR) repeats.

SO Sequence 397 BP; 124 A; 104 C; 104 G; 65 T; 0 other;

Query Match 1.5%; Score 56.2; DB 20; Length 397;  
 Best Local Similarity 47.3%; Pred. No. 0.0019;  
 Matches 169; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 1454 cagcaccacaaatgttgcacacgcctcctactcttagtggtgtacaacttactactgtgt 1513  
 Db 358 cccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 299  
 QY 1514 cctctacaacaatgttggggaatgtgctggtgtgtaagtgaactggaactatt 1573  
 Db 298 ctg 239  
 QY 1574 atccctggtctcaattacacacacgcttgcctactaagtgtaagcgaacttactttagctg 1633  
 Db 238 ctg 179  
 QY 1634 ataacttgcattatagtgtaaaaagaatcgtgtacagttgtgtcttcgaactaataata 1693  
 Db 178 ctg 119  
 QY 1694 atgcagacgctaaatgttaactttagtgcacaaagcgaccccttaagaaatacaatct 1753  
 Db 118 tcgaaacatttcaaaactgaagttatttttttttttttttttttttttttttttttttttt 59  
 QY 1754 atcaaatcaaatatnaaagcgcttggaagtcgataaggtlatgttaagtaggcactg 1810  
 Db 58 atgagaaaactatatttcataagaaaataacattgtttcaccgaatcaagtagtctactg 2

RESULT 12

T42063/C  
 ID T42063 standard; DNA; 1830121 BP.

AC T42063;

DT 14-SEP-1999 (first entry)

DE Haemophilus influenzae complete genome sequence.

KW Genome: bacterium; Haemophilus influenzae; computer readable medium;  
 KW expression modulating fragment; regulation; gene expression; vector;  
 KW organism; open reading frame; ORF; ds.

OS Haemophilus influenzae.

PN MO9633276-A1.

PD 24-OCT-1996.

PF 22-APR-1996; 96MO-US05320.

PR 07-JUN-1995; 95US-0487429.

PR 21-APR-1995; 95US-0426787.

PR 07-JUN-1995; 95US-0476102.

PA (HUMA-) HUMAN GENOME SCI INC.

PI (UJJO) UNIV JOHN HOPKINS.

PI Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;  
 DR WPI; 1996-485782/48.

PT Haemophilus influenzae Rd genome recorded on computer readable  
 PT medium - useful for identifying commercially important nucleic acid  
 PT fragments by homology searching

PS Claim 1; Page 77.2-77.1091; 1291pp; English.

XX This sequence represents the complete genome sequence of the bacterium  
 CC Haemophilus influenzae strain Rd. The invention relates to a computer  
 CC readable medium (CRM) having recorded upon it the complete H. influenzae  
 CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide  
 CC sequence at least 9% identical to (I). By providing the full-length  
 CC genomic sequence in a computer readable form, it is possible to identify  
 CC commercially important nucleic acid fragments and expression modulating  
 CC fragments (EMPs) of the Haemophilus genome. The EMFs can be used to  
 CC regulate the expression of a nucleic acid molecule. Vectors and altered  
 CC organisms comprising the predicted ORFs can be used to produce any of the  
 CC polypeptide fragments of the H. influenzae Rd genome.

SO Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Query Match 1.4%; Score 53; DB 17; Length 1830121;  
 Best Local Similarity 45.5%; Pred. No. 0.12;  
 Matches 277; Conservative 0; Mismatches 320; Indels 12; Gaps 2;

QY 601 gatattaccgttcacatgaagcgcggaagctacgaagcgtgtgaaagcagcaattta 650  
 Db 1806740 ggtagcagcaggtgacaaacccctgacgttaacaaagatggatttaaacaggtatataa 1806691  
 QY 661 gaaacagctaaagctgtgtgtgaatcaatcaatcaatcaatcaatcaatcaatcaatca 720  
 Db 1806680 gctattacttaattggtgaggtggttaagagctttgacgatgcaatttt--tgatggt 1806624  
 QY 721 ttcaaacctgaactaagcagaagcagcaaaagccttgatgacatgaagctgctgt 780  
 Db 1806623 tttaatttaactgcaacgatttttaattgacacggttggaagatgcttataaagttattta 1806564  
 QY 781 atcccaaaagcttgaagctgaaagtgatgaacacacacaaagcctgttgattacca 840  
 Db 1806563 aattcttaattgaaaaaaatgcaaat-----aaacaaacgcttggtgcttacgacacg 1806513  
 QY 841 gcagtaaccagtgaaagcgaacactaaatcaacttccagctgtgtgcaaatgaagataca 900  
 Db 1806512 gcgggagactgtaggggatttactgtaattggttggtggtatgacaaacaaacggtacg 1806453  
 QY 901 gtaaacgttaatactgtacgtatcataaagtgacggtgaacattccattgcccattat 960  
 Db 1806452 aaaaagaaagcaatcaagtttaaacagctgataagtccttttaacgacccggtgct 1806393  
 QY 961 acgacagatgttcttcttactacagcgaagaaactacactacactgtgtgacttcaactca 1020  
 Db 1806392 gcttagcggttacttccaaatcttaaaacggttaacattacgatttaccggtttatgta 1806333  
 QY 1021 ttcaaaataatacggagatataaagtagtataaagtagtataaagtagtataaagtagt 1080  
 Db 1806332 acttaaaagcgaatgagcgttcttaaaaaagatggcgatattatgacgtcaaaatgatatat 1806273  
 QY 1081 gaatttaagaagatgcatcattcactcaagctcgaatgagatgctgtgaattgactcaagtg 1140  
 Db 1806272 caaaacactgattatggttttaactgttggttaattggtgactgctgacacttaagtgac 1806213  
 QY 1141 ttggaactaagttaacaacaacacacttctgttaacttagcagcaggtgacttgcagact 1200  
 Db 1806212 tttagaaactgtttaaactgagcagcactgacagatccggttaaggtatgtaaaaght 1806153  
 QY 1201 gacgatact i209  
 Db 1806152 gcttagctgt .1806144

RESULT 13

X61812  
 ID X61812 standard; DNA; 1039 BP.

```
XX AC X61812;
XX DT 19-JUL-1999 (first entry)
XX DE B. burgdorferi antigenic protein coding sequence, t5-14.nt.
XX KM Antigenic protein; vaccine; Lyme disease; infection; detection: ss.
XX OS Borrelia burgdorferi.
XX PN W09859071-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US12718.
XX PR 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MEDI-) MEDIMUNE INC.
XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX DR WPI; 1999-189980/16.
XX DR P-PSDB; Y20115.
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
XX PT products for the diagnosis, prevention and treatment of diseases
XX PT caused by Borrelia, particularly Lyme disease
XX PS Claim 1; Page 202; 275pp; English.
XX CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
XX CC invention, which is suitable for use in a vaccine. The Bb polypeptides
XX CC can be used in vaccines for eliciting protective antibodies to members of
XX CC the Borrelia genus, particularly for the use against Lyme disease in
XX CC humans and animals. They can be used for preventing or attenuating an
XX CC infection caused by a member of the Borrelia genus. The products can also
XX CC be used for detection of members of the Borrelia genus.
XX SO Sequence 1039 BP; 526 A; 94 C; 212 G; 207 T; 0 other;

Query Match 1.4%; Score 51.8; DB 20; Length 1039;
Best Local Similarity 47.6%; Pred. No. 0.023;
Matches 152; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 177 tggtagattcccaacttaacgtatgtatgtcgtgaatacacaaagcgaanaacgata 236
DB ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 237 ccgtgacgcgtgacattagtaataagcagtgcgcgaaanaagcgttacttagc 296
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 297 tgaattcaaaagaatataaacttcglttcaaaagcaaaccttaattcgtgcgaagc 356
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 tcaaaagcaaaaagaagaagaaactctaaagaagaacacaaacaaagaagaagc 459
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 aaaaagctagagcagatgaaagaagaagacgaactaaacaaacaaagaagaaga 519
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 agagctagagcgtcgtcttcaagcaaaagatttagaaanaagcagaataactatcaca 476
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 acagggaagaagctaggggttaagcagaanaaagaagaagaagaagaagaagaaga 579
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 aattccttaagaaataa 495
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
DB 580 acaagaagaagaagaaga 598

RESULT 14
X61811
ID X61811 standard; DNA; 1125 BP.
XX X61811;
XX DT 19-JUL-1999 (first entry)
XX DE B. burgdorferi antigenic protein coding sequence, t5-14.nt.
XX KM Antigenic protein; vaccine; Lyme disease; infection; detection: ss.
XX OS Borrelia burgdorferi.
XX PN W09859071-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US12718.
XX PR 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MEDI-) MEDIMUNE INC.
XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX DR WPI; 1999-189980/16.
XX DR P-PSDB; Y20114.
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
XX PT products for the diagnosis, prevention and treatment of diseases
XX PT caused by Borrelia, particularly Lyme disease
XX PS Claim 1; Page 202; 275pp; English.
XX CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
XX CC invention, which is suitable for use in a vaccine. The Bb polypeptides
XX CC can be used in vaccines for eliciting protective antibodies to members of
XX CC the Borrelia genus, particularly for the use against Lyme disease in
XX CC humans and animals. They can be used for preventing or attenuating an
XX CC infection caused by a member of the Borrelia genus. The products can also
XX CC be used for detection of members of the Borrelia genus.
XX SO Sequence 1125 BP; 565 A; 101 C; 223 G; 236 T; 0 other;

Query Match 1.4%; Score 51.8; DB 20; Length 1125;
Best Local Similarity 47.6%; Pred. No. 0.024;
Matches 152; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 177 tggtagattcccaacttaacgtatgtatgtcgtgaatacacaaagcgaanaacgata 236
DB ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 237 ccgtgacgcgtgacattagtaataagcagtgcgcgaaanaagcgttacttagc 296
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 297 tgaattcaaaagaatataaacttcglttcaaaagcaaaccttaattcgtgcgaagc 356
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 tcaaaagcaaaaagaagaagaaactctaaagaagaacacaaacaaagaagaaga 422
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 agttagagaggaagaagaagaacgaagtagtaanaaagaagaagaagtggagaagcaag 482
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 297 tgaattcaaaagaatataaacttcglttcaaaagcaaaccttaattcgtgcgaagc 356
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 483 tcaaaagcaaaaagaagaagaaactctaaagaagaacacaaacaaagaagaagc 542
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 tctgttagcaacttacatcgatgttcaacaactatgcaacaanaattagagcaattgcgca 416
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 543 aaaaagctagagcagatagagaagaagaagacgaactaaacacacaaagaagaagaaga 602
DB || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



